DOCUMENT RESUME

ED 174 650 TM 009 499

AUTHOR	Airasian, Peter W.; And Others
TITLE	Policy Issues in Minimal Competency Testing and a Comparison of Implementation Models. Report Submitted to the Policy Subcommittee of the Massachusetts Advisory Committee on High School Graduation Requirements.
INSTITUTION	Heuristics, Inc., Dedham, Mass.
SPONS AGENCY	Massachusetts State Dept. of Education, Boston.
PUB DATE	14 Apr 78
NOTE	161p.
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Academic Achievement; Academic Standards; Basic Skills; *Educational Assessment; Educational Policy; Elementary Secondary Education; Estimated Costs; Evaluation Needs; Graduation Requirements; Legal Problems; *Minimum Competency Testing; Models; *Program Development; Senior High Schools; State Programs; *State School District Relationship; State Standards; *Testing Problems; *Testing Programs

ABSTRACT

The justification of minimal competency programs relative to perceived educational problems, the focus on individual pupil certification and sanctioning, and the lack of clear resolution on many technical problems, give rise to a number of policy issues critical to minimal competency testing programs which are tied to graduation. Three program models are described and compared across seven of these policy issues: strong state control; large impact on curriculum; increased remediation; difficulty in setting standards; measurement problems; legal challenges; and high costs. The models vary according to state vs. local school district responsibility for selecting competencies, tests, and standards. All policy issues are so complex that it is not possible to recommend a particular model, or to decide whether a testing program should even be implemented. However, several recommendations can be made regarding the adoption of minimal competency testing. The general public should be made aware of its impact on the individual student. If a program is implemented, an adequate phase-in time should be allowed for curriculum revision and test validation. Pilot programs should be implemented and compared--no program need be viewed as the final one. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

ED174650

Commonwealth of Massachusetts

Department of Education

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
C. DiBella
MA. SDE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Policy Issues in Minimal Competency Testing and a Comparison of Implementation Models

TM009 499

HEURISTICS INC.

Policy Issues in Minimal Competency Testing and a Comparison of Implementation Models

Report Submitted to the Policy Subcommittee of the Massachusetts Advisory Committee on High School Graduation Requirements

Prepared by:

Peter W. Airasian
George F. Madaus
Joseph J. Pedulla

with the assistance of

Kenneth B. Newton

April 14, 1978

TABLE OF CONTENTS

	Page
PART I: SUMMARY	1
PART II: OVERVIEW OF THE ISSUES	10
Introduction	10
General Issues Associated with Minimal Competency Programs	16
1. What is meant by the term minimal competency?	17
2. Who is held accountable in minimal competency graduation programs?	22
3. What characteristics should the selected competencies possess?	27
4. When should pupils be tested for minimal competence?	34
5. How should special pupil populations be dealt with?	38
6. How much need will there be for remedial programs?	41
7. Will Minimal Competency Tests impact on the curriculum?	48
8. How are standards for passing a minimal competency test to be established?	54
9. What are the problems of test security in minimal competency testing programs?	63
10. What are the issues of test validity?	66
11. What are the issues related to the difficulty level of items on a minimal competency test?	74
12. Can competency and item pools be developed?	76
13. What are the costs of a minimal competency testing program?	77
Possible Costs	79
Factors Influencing the Magnitude of Costs	84
Cost Estimates	88
Summary	90
PART III: COMPARISON OF SELECTED MODELS	94
Model 1: State responsibility for specifying minimal competencies, establishing testing procedures and setting standards	97
Issues of State versus Local Control	101
Impact on Curriculum	103
Remediation	105

TABLE OF CONTENTS (cont.)

	Page
Standards	106
Measurement Issues	108
Legal Issues	111
Costs	112
Model 2: State provided options for competencies, testing procedures and standards from which local districts select	115
State versus Local Control	121
Impact on Curriculum	122
Remediation	124
Standards	125
Measurement Issues	127
Legal Issues	128
Costs	129
Model 3A: Local responsibility for setting competencies, testing procedures and setting standards subject to state approval	
Model 3B: State responsibility for setting competencies, local responsibility for testing procedures and setting standards subject to state approval	131
State versus Local Control	136
Impact on Curriculum	137
Remediation	138
Standards	139
Measurement Issues	140
Legal Issues	142
Costs	143
Summary	145
PART IV: RECOMMENDATIONS FOR IMPLEMENTATION	147
BIBLIOGRAPHY	154

PART I: SUMMARY

The purpose of this report was to examine policy issues associated with minimal competency testing programs which are tied to high school graduation. A large number of general policy issues were identified, discussed, applied to three specific implementation models or strategies, and finally recommendations for implementation were made. The body of the report is necessarily long, containing a myriad of policy concerns pertinent to minimal competency programs tied to high school graduation. Questions of program need, definition, impact on curriculum, provision of remedial services, measurement problems and cost are presented. To judge from practices currently being carried out in other states and from the rhetoric surrounding minimal competency testing, one might conclude that a report like this is unnecessary, that minimal competency programs are simple to conceive and carry out. Our report makes clear that this conclusion is unwarranted. We shall not attempt to summarize every point and issue raised in this report; instead the reader is urged to read the full report. Our purposes in this summary are four: 1) to provide a general context within which policy issues related to minimal competency programs may be placed and understood, 2) to highlight some of the more pertinent policy issues we have identified, 3) to discuss the implementation models, and 4) to summarize our recommendations.

There are three main points which one must bear in mind in thinking about minimal competency programs in which pupils' performance is tied to high school graduation. These points provide the general context within which minimal competency programs must be addressed. First, we must ask ourselves, "Why

minimal competency testing programs; what needs do such programs respond to?" There are many reasons given to justify minimal competency testing programs, including: increasing pupil achievement, making schools more accountable, making the high school diploma more meaningful, providing early diagnosis and remediation to poor learners, identifying poor teachers or schools in terms of poor student performance on the basic skills, improving the cost effectiveness of education, and so on through a veritable litany of presumed benefits. As our report points out, there are legitimate questions as to whether minimal competency testing programs per se respond to these perceived needs. For example, can we increase pupil achievement by legislative fiat? Do we have the knowledge and understanding necessary to teach all pupils basic skills? Is the idea of beginning minimal competency testing at grade 8 antithetical to early diagnosis and remediation for poor school learners? Will increased costs resulting from a minimal competency program erode any potential benefits it appears to hold out? Will minimal competency test results permit us to hold schools or teachers accountable for pupil learning; will these tests provide the type of accountability information we seek? These and a series of other questions related to the perceived ends of minimal competency testing are explored in our report. Suffice it to say here that it is not at all evident that a minimal competency testing program can respond to all these perceived needs. We have urged that the primary needs people perceive being met by minimal competency programs be articulated and that these needs be examined in the light of whether such programs, as currently conceived, actually respond to these needs. In many cases minimal competency testing programs

will be found lacking. Minimal competency testing is far from being a panacea for educational ills - be they real or imagined.

A second, and exceptionally critical point, which must be kept in mind when considering the policy implications of minimal competency testing programs which tie pupil performance to receipt of a high school diploma, is that such testing programs are vastly different in process and implications from most other testing programs with which we have had experience. Minimal competency testing is not just another testing program; it is different in kind and consequence from existing programs.

In most existing testing programs, be these in the form of state-wide assessments or local district achievement testing, inferences about student learning, program effectiveness and the like are made on the basis of aggregated data, or more simply data which are summed and averaged over a large number of pupils. Decisions affect individual pupils only insofar as they affect the pupils instructional program or activities. The consequences associated with individual pupil performance are not generally great or severe. Minimal competency testing programs in which pupil performance is tied to high school graduation are not primarily interested in the average performance of groups of pupils, but rather in the particular performance of <u>each individual</u>. Moreover, the consequences of poor performance are grave for <u>each</u> pupil, namely denial of a high school diploma. Because the focus in minimal competency testing is on <u>each individual</u> pupil, and because there are severe penalties associated with poor performance, minimal competency programs are unlike most existing testing programs.

The uniqueness of minimal competency testing is not an argument that such testing cannot or should not be carried out nor that receipt of a high school diploma should not be based upon such testing. However, the unique problems posed by this type of testing suggest strongly that it must be carried out very carefully. Tests must be validated, standards shown to be unbiased, remediation provided where necessary, etc., if the fairness of denying a single pupil a diploma based on the minimal competency test is to be defended.

In discussions about minimal competency programs, one hears more and more about legal challenges and issues which have arisen in other states. The primary basis for such challenges is injury to an individual pupil, usually denial of a high school diploma. The challenges are directed towards specific features of existing minimal competency programs such as the fairness of testing procedures vis a vis instruction, the bias toward certain racial, ethnic and special needs populations, the failure to provide necessary remedial services, lack of sufficient phase in time, etc. Notice, that in state-wide assessment, or local district achievement test programs, such issues are not the bases for legal challenge. The difference in minimal competency programs is that decisions are being made about, and sanctions are being applied against, individual pupils; this is not the case in our usual testing programs. It is for this reason that legal concerns with due process and injury are arising in the context of minimal competency programs.

Such threats and, of course the ethical concern of waiting to institute a rigorous but fair program, obviously give rise to a number of policy issues regarding the implementation of competency testing programs. For example, test validation and standard setting become critical policy issues in implementing programs.

Test administration and scoring, the provision of remedial services, and the opportunity pupils have to learn the desired competencies, as well as the cost of insuring that these activities are carried out, also assume importance as policy issues. The report points out concerns associated with these and other policy issues as they apply to the problem of certifying individual pupil competency as opposed to surveying group achievement.

The third policy referent is related to the second. As just noted, certification of individuals, with potentially important sanctions associated with lack of certification, lead to the necessity of addressing a number of technical issues related to testing procedures and standard setting. For example, how will issues of potential test bias be dealt with? How can we insure that test topics and items measure school learning and not more general, non-school factors such as social class, home background, or intelligence? This issue is important, since if we are to certify pupils individually, we want to be certain that they have had an opportunity to learn the measured competencies as part of their school experiences. When should tests be direct measures of skills and when will indirect measures suffice? Must testing procedures be timed or untimed and are open-ended test situations necessary for some skills? How does one set valid, defensible and unbiased standards of minimal competency?

These, and other policy questions arise in the context of minimal competency programs because of their emphasis on individual pupil certification. The difficulty is, however, that in many instances there are no well-defined or commonly accepted answers to many of these questions. Setting standards of performance for minimal

competence is problematical; simple, arbitrary solutions to this problem tend to result, as in Florida, in legal action, and a loss of program credibility. Overall, the technical issues associated with the implementation of minimal competency testing programs are not simple and easily resolvable, as is pointed out in this report.

These three points, 1) the justification of minimal competency programs relative to perceived educational problems, 2) the focus on individual pupil certification and sanctioning, and 3) the lack of clear resolution on many technical problems, give rise to a number of policy issues in conjunction with the installation of minimal competency testing programs tied to graduation. Too often, the focus on the perceived ends of such programs have blinded individuals to these points and issues. The major points made in Part II of the report, the overview of the issues, are as follows:

1. Various publics (parents, teachers, employers, representatives of the poor and minority groups, etc.) should be actively involved in stating or selecting the minimal competencies, and these competencies should be justified in terms of these publics' consensus on their importance and not in terms of their being essential life skills or necessary for adult functioning.

2. Most minimal competency programs are structured in such a way that the pupil is the only individual who can possibly be held directly accountable for failure to demonstrate the minimal competencies. By testing from grade 8 onward, poor performance can only be attributed to the pupil or the "system". It would be quite unfair to try to hold an individual teacher accountable for a student's failing the minimal competency test.

3. Minimal competency programs are apt to impact most heavily on that segment of the student population which is weakest scholastically. These students should be identified as early as possible, certainly before grade 8, so that remedia ion can begin. One should be aware that this early identification could pejoratively label these students and set in motion a self-fulfilling prophecy of failure.

4. For a competency to be included in a minimal competency testing program, there should be a strong belief that the competency can be taught to, learned by, and, where necessary, remediated for all students. This remedial component must be an integral part of any minimal competency program.

5. The number of competencies selected should be kept small and should be drawn from basic cognitive areas that are presently part of the school curriculum. Affective or adult functioning competencies should not be included as part of a minimal competency program tied to high school graduation.

6. Testing should occur early and often so that multiple opportunities to sit for the test are afforded to pupils. Once a student has passed the test, he/she should not be required to demonstrate the competencies again at a later date in order to graduate.

7. Most handicapped pupils should be treated as individual cases in determining how, of if, they should be required to demonstrate the required competencies.

8. A minimal competency program may not result in an increase in remedial services over present levels, but may re-focus those services, in terms of record keeping and testing, in order that a school can document that any student in need has received the required services.

9. The effects of a minimal competency program on the school curriculum are a function of the perceived importance of the test, the extent to which the competencies are already a part of the curriculum, and the amount of lead time before the required program takes effect.

10. It is not possible nor desirable to try to use the same test year after year. The tests should be secure only up to the time they are administered and then should be made available for public scrutiny.

11. Setting a cut-score, or standard, be it norm- or criterion-referenced, is always a value-laden, judgmental process. There are no agreed upon statistical procedures that remove human judgment from setting standards.

12. One must be able to demonstrate that the minimal competency test samples content and behaviors which are part of the school's curriculum. Demonstrating this fact could require extensive record keeping for each pupil.

13. There are a wide range of direct and indirect costs associated with the development and maintenance of a minimal competency testing program which must be clearly anticipated in advance so that adequate funding can be provided. Direct performance measures, which are more costly than indirect, paper and pencil measures, may be necessary in testing some competencies.

Part III of the report compared three potential implementation models: 1) the state specifies the competencies, develops tests, and sets cut-scores for passing the tests; 2) each local district selects a package from a group of packages of competencies, tests and cut-scores developed by the state; and 3) local districts specify their own competencies, develop their own tests, and establish their own cut-scores with all three aspects subject to state approval. A variation on this third model was also included for comparison, with the one difference being that the state sets the competencies. These models were then discussed in general and compared on seven criteria: 1) state vs. local control, 2) impact on curriculum, 3) remediation, 4) setting standards, 5) measurement issues, 6) legal issues, and 7) costs. Because the issues are very complex and interrelated, and because there are trade-offs associated with each model, it is not possible to provide a brief summary of Part III. Suffice it to say that each model had its strengths and weaknesses, and no one model emerged as clearly preferable to the others.

The final section of the report provided recommendations on how best to proceed in determining whether or not to implement a minimal competency

testing program tied to high school graduation. These recommendations were of a general nature, suggesting information and data that should be gathered prior to implementing any minimal competency program. It was also recommended that pilot programs, using different implementation models, be instituted to answer questions that simply can not be answered at this point.

We suggest that the resolution of the many policy dilemmas resulting from consideration of a minimal competency testing program is not to abandon all plans for such a program. Rather, the lesson to be learned from the points raised in this report is to proceed slowly. Minimal competency testing programs may hold promise, but in the present climate it is fair to say that such programs present enormously complicated problems. Before performance on minimal competency exams is tied to high school graduation, care must be taken that the conceptualization, procedures, and implementation of the program is as fair and equitable as possible. If we are to place the burden of proof for learning on students, it is only fair that we place the burden for a fair, equitable and well thought out minimal competency program on ourselves.

PART II: OVERVIEW OF THE ISSUES

Introduction

Presently thirty-two states have enacted or have under consideration some form of minimal competency testing. This movement toward minimal competency certification is the most recent, and perhaps the most significant, manifestation of the accountability movement in education which has been with us for the past decade or more. Given the large number of states considering some form of a minimal competency program, it is not surprising that the term means many things to many people. For some minimal competency testing (MCT) is viewed as a certification mechanism whereby a pupil must demonstrate that he/she has mastered certain minimal skills in order to receive a high school diploma. This report examines in detail this particular conceptualization of MCT. Others see MCT simply as a device to increase the emphasis schools place on the three R's or basics. Still others see MCT variously as a mechanism for tightening up promotion requirements; certifying early exit from the school system; holding educators responsible for poor student achievement; increasing the cost-effectiveness of education; identifying and remediating pupils who are having learning difficulties; or increasing the publics' confidence in the schools and their graduates.

Regardless of how MCT is conceptualized, a common, implicit assumption underlying all viewpoints, is that pupils should, and in fact can, master the skills associated with the three R's. While a number of MCT programs extend beyond the three R's to include life, survival or consumer skills and competencies, emphasis on the three R's underlies all approaches.

We should all like to be assured that, after 12 years of schooling, pupils possess, at a minimum, basic skills in the areas of reading, writing and computation. That such a goal is politically, economically and educationally desirable (although not necessarily in that order) few would debate. However, we may contrast the problem of low pupil achievement with that of equal educational opportunity, defined as equal access to curricula, schools or facilities. The latter problem is essentially a political one and can be rectified largely by legislative action - as indeed has been the case in recent years. However, that pupil achievement is manipulable or can be affected directly by legislative fiat is an unresolved issue. The central problems of minimal competency programs, especially those in which measured competence is tied to high school graduation, are complex in nature. They involve issues such as "what is minimal competency and how is it defined?"; "are presently existing tests suitable for minimal competency certification?"; "how does one set fair and defensible standards?"; "can competencies be taught to all pupils with reasonable hope of success?"; "are suitable and effective remedial strategies available to help hard-core unachievers master the competencies?"; etc. These are questions of means, not ends, and it seems clear that some states have overlooked the means problems or assumed that they are readily resolvable. One of the primary functions of this document is to consider such means issues as they relate to the educational, measurement, legal and cost implications of minimal competency programs.

Massachusetts pupils will be the ultimate beneficiaries of the fact that the Commonwealth, through the Bureau of Research and Assessment, has chosen to enter the minimal competency arena relatively late, and in a way that seeks to identify as many issues and potential problems as possible before making a firm commitment to establishing a minimal competency program. State authorities are thus in a position to learn from the experience of other states and to avoid many difficulties which have been encountered in states which have enacted MCT graduation programs. Certainly any program which might be instituted should be carried out in a way that is equitable and defensible to students, parents, teachers, and school administrators. In this context, it seems more reasonable to understand the potential problems associated with minimal competency programs - even if some of these problems cannot be resolved unequivocally - than to launch a program blind to its possible human and financial costs.

By way of background it should be pointed out that this document is the result of three months work on the part of the authors and the six member Policy Subcommittee of the Massachusetts Advisory Committee on High School Graduation Requirements. The authors were charged to carry out four tasks: 1) identify a series of models under which a minimal competency program beginning at grade 8 and tied to high school graduation could be carried out; 2) identify and discuss areas of policy concern associated with MCT programs (e.g., legal, measurement, cost, etc.); 3) compare three minimal competency models in terms of the identified areas of concern; 4) make suggestions regarding

implementation strategies and research activities which should be carried out regardless of which minimal competency model might be adopted.

The remainder of this report is concerned with these four areas. What follows in this section is a general overview of the many issues underlying and associated with minimal competency programs. This section considers the philosophical, definitional, educational, measurement, legal and cost issues pertaining to minimal competency programs. Part III focuses on the three minimal competency models selected by the Policy Subcommittee for detailed investigation. A description of how these models were derived as well as a comparison of these models on the basis of educational, legal, measurement and cost criteria is presented. The final section of the report contains general recommendations for implementing whatever minimal competency model might be selected by the Commonwealth.

Before turning to general issues associated with minimal competency testing, a brief discussion of why MCT has suddenly burst on the educational scene with such considerable grass roots support is in order. There are a number of interrelated reasons to explain the rapid growth of the MCT movement. First, starting in the mid-1960's, a number of large scale studies of schools and compensatory education (e.g., Coleman et al, 1966; Cicirelli et al, 1969; Jencks et al, 1972) concluded that schools and school programs had very little direct, or unique effect on the achievement of pupils. Moreover, these studies implied that educational expenditures in the form of small class sizes, teacher salaries, and school resources had little impact on the achievement of pupils.

While these conclusions have been hotly debated, it is clear that they were much publicized and led many people to rethink their beliefs about the efficacy of schools and school programs.

More recently, the 14 year decline in national SAT scores has been discovered and publicized as indicative of a progressive decline in pupils' basic skills. Despite the fact that the SAT skills are not a measure of basic skills many have interpreted the SAT score decline in this simplistic way. Others have interpreted the SAT score decline as indicative of a general decline in "standards". In addition to the widely publicized SAT score decline other investigators also have noted a decline in achievement scores associated with state testing programs and some have attributed these drops to a lowering of "standards", a falling off in the mastery of the three R's.

In 1972 a San Francisco public high school graduate filed the first "educational malpractice" suit arguing that he had been graduated without being taught basic skills. Although Peter Doe lost his case, the question of the competence of high school graduates in general was implicitly raised. Moreover, the threat of similar law suits has made many educators more cognizant of the need for instruction and certification in basic skills.

More recently, employers and college teachers have joined the litany of voices decrying the lack of reading, writing, and mathematical ability of high school graduates. The voice of these groups has been quite influential in media accounts of how the schools are failing or how "standards" are falling off.

We may add two additional explanations for the near band wagon emphasis of MCT programs focusing on the three R's and "basic skills". First, as just

noted, there is a band wagon in motion, and it is increasingly difficult for school persons and legislators to ignore it or to avoid being swept up in its movement. Second, and at a much more general level, the call for a return to basics may reflect a societal malaise resulting from the false expectations of the Great Society programs and the recent past history of an emphasis on openness, "relevancy" and electives in school curricula. In large part the programs of the Great Society, for a variety of reasons, placed educational reform in a central position. Expectations about the power of education to make all pupils literate as well as to solve wider social and economic ills were first raised, and subsequently dashed, over a period of ten years. The increased emphasis on basic skills may reflect a retrenchment in the expectations for education as a force for social and economic change, and a concern that educators concentrate on those areas in which most people perceive there was success in the past, namely, teaching the three R's. If education cannot reform the whole society, at least it can teach pupils to read, write and compute.

Further, it is clear that the back to basics cry cannot be separated from a growing disenchantment with the curricula innovations of the late '60's and early '70's. Whether the view is correct or not, there is a feeling abroad that innovative curricular changes resulted in a de-emphasis on the basic skills. Those holding this view point to the test score declines as clear evidence that their argument is correct.

These, then, are the reasons behind the strong support MCT programs have received in many quarters. Although it is not evident which of these many reasons is the most influential in the move toward MCT programs, it seems clear that they all have contributed to some extent. The accuracy or validity of these perceptions about the effectiveness of schooling can certainly be questioned. However the accuracy of such perceptions is not the central issue. The central issue is that these perceptions exist and persist and because of this fact there is a growing sentiment that something needs to be done, that there is a need to restore standards and to move back to "basics". For many, MCT appears to be the mechanism for obtaining these goals. The policy-maker needs to be aware of this mind set when considering MCT.

GENERAL ISSUES ASSOCIATED WITH MINIMAL COMPETENCY PROGRAMS

As one examines the underlying rationale and mechanics of existing minimal competency programs, a number of questions and concerns inevitably come to mind. In this section we shall endeavor to point out these questions and concerns, not so much to resolve them (many are, as we will see, unresolvable except by administrative or legislative fiat), but to identify and provide an overview of them. While many of the questions and concerns pertain directly to the implementation of minimal competency programs (e.g., costs, testing procedures, standard setting) others are questions of values or terminology. (e.g., the nature of the competencies sought, the language of minimal competency). The latter are included in our discussion because they can impact upon imple-

mentation decisions and because we feel that it is essential that these kinds of issues be explicated if informed decisions about the desirability and nature of MCT programs are to be made. This section, then, presents an overview of a number of questions and concerns which are of import in considering any MCT program which is tied to the awarding of a high school diploma.

1. What is meant by the term minimal competency?

Although this question is posed in terms of the meaning of the phrase "minimal competency", the issue is more general, dealing with the language of competency programs and the effect of the language on the programs' appeal. We noted above that the language of competency tends to be ends oriented, focusing upon the potential products of programs rather than upon the means necessary to attain these ends. An emphasis on ends or outcomes is not unexpected, since much of the rhetoric about minimal competency programs has been generated by advocates seeking to propagate their belief in the efficacy of such programs. However, much of the appeal of minimal competency is a function of its stated ends, rather than a function of any evidence of the ability of such programs to attain these ends. Further, the language employed in discussions of minimal competency programs, unfortunately, has the potential of raising expectations about the ability of schools to make all segments of the student population literate, without the certainty that schools can actually attain this end.

Consider, for example, the language used by the California State Legislature when it passed a law requiring junior and senior high schools to adopt standards of proficiency in basic skills and tying high school graduation to these standards by 1980: "pupils attending public schools in California acquire the knowledge, skills, and confidence required to function effectively in contemporary society." Such language is not unique to the California Legislature and a similar theme can be found in bills of other states. At first blush, it is difficult to fault such an intent. It holds out the promise of knowledgeable, confident pupils, successfully carrying out their chosen roles in society. Few would debate the merits of such a goal.

However, if we consider more carefully the words used in the intent we shall see that there are a number of questions left unanswered by the legislature's statement. These questions are important since, in two years, receipt of a high school diploma in California will depend upon their answers. Consider the wording more closely. What, for example, is the distinction between knowledge and skills? Is knowledge a familiarity with basic facts such as the multiplication tables, rules of punctuation, etc., while a skill is an application of these facts to real life situations such as filling out a tax form, writing a coherent business letter, etc.? And if this distinction is correct, are we confident that schools can - or should - teach such applications to all pupils? Should every pupil have to be able to fill out a tax form, compute compound interest, etc., in order to receive a high school diploma? (A poll of the three authors of this report indicates that none has filled out his own tax form in recent years.)

More to the point, however, are the questions raised by the words used in the latter portion of the statement. Can - or should - schools teach confidence? What exactly is confidence? How much confidence is too much and how much is too little? Is the possibility that some pupils may fail to demonstrate the cognitive proficiencies required for obtaining a high school diploma antithetical to the goal of instilling confidence in all pupils? What of the pupil who can demonstrate cognitive competencies but lacks confidence? Should he/she be denied a diploma? At a more general level, the intent to instill an affective behavior such as "confidence", in all pupils raises a number of legal, ethical and political questions about the aim and function of schools in the affective domain.

Finally, what skills, knowledge, etc., are "required" to "function effectively" in "contemporary society"? There is a legitimate question - and this question should be examined empirically - as to the number of American adults who presently possess the basic skills and knowledge talked about in many minimal competency programs. What then, does "function effectively" mean? Is there a universal norm of effective functioning which can be applied to all people, regardless of occupation, geographic location, level of education, etc., to define the ability to function effectively? Moreover, in a society which is changing constantly, is it sufficient to set our goal at effective functioning in "contemporary society"? "Contemporary society", by definition, will not be contemporary in five or ten years and the skills, knowledge and confidence

required to function in 1978 may be quite obsolete by 1988. Consider such innovations as the advent of the hand calculator, the installation of cash registers which compute change, television journalism, and the like and it is clear that many of the skills which were held to be critical to adult functioning in the recent past are in the process of becoming obsolete - or at least dispensable.

We have singled out the California legislation for discussion, but we might have made the same arguments using other states as examples. Had we done so, the arguments made would have been similar, but the terms would have been different (e.g., survival skills, minimal competence, life skills, functioning members of society, etc.).

If one examines the California State Legislature's intent vis a vis its minimal competency program, one sees that the second half of the intent "required to function effectively in contemporary society" is really a justification for the first half "acquire the knowledge, skills and confidence". We have just seen that the fuzziness of this terminology and the questions it raises cloud both the intent and specifics of the legislation. It appears that there are at least two ways to justify the need for a minimal competency program. The first is to justify the program in terms of some wider societal functioning and adequacy, which is what the California legislators have done. The second is to say, simply, that high school graduates should have certain competencies because the community believes and wants pupils to have them and avoid statements which directly tie them to "survival" or "effective functioning" in society.

As we shall see, the way in which the legislature answers this question directly impacts on the type of validity a test of minimal competency must possess.

It needs to be understood that the adoption of a particular MC program and its rationale reside in the value domain. Whether or not we will have a minimal competency program in Massachusetts will not be decided on the basis of empirical evidence. Such evidence may explain, or describe, particular characteristics of the MCT program, but the evidence itself will not determine its adoption or defeat. People will do this, and they will do it on the basis of value judgments about the worth of the program and of the competencies such a program encompasses.

We do not wish to prolong this discussion about the language of competency, but it is worthwhile to point out the seductive nature of the terminology used to describe and justify minimal competency programs. Undefined, perhaps undefinable terms are used without consideration in discussing MCT programs, and it is only when one thinks through the meaning and application of such terms that the apparent simplicity of MCT is stripped away revealing its true complexity. The present language of MCT programs often does a grave disservice by masking this complexity which in turn falsely raises expectations. It will serve the policy committee and the legislature well to think carefully about the clarity of the language used in connection with legislation regarding a MCT program tied to graduation.

2. Who is held accountable in minimal competency graduation programs?

We noted above that the minimal competency movement is the latest manifestation of the more general "accountability movement". One central issue then, is in a given MCT program, who is ultimately accountable for student performance? Although different publics wish to hold teachers or parents or the schools accountable for students' poor achievement, it is clear that when minimal competency testing is tied to high school graduation, the onus for failing falls squarely on the pupil.

There are many reasons that the primary impact falls on the pupil, ranging from the failure of states and locales to consider in advance of program implementation just who they wished to hold accountable, to political and logistical problems in trying to hold anyone save students accountable for their own learning. Given that much of education is sequential and cumulative and that most skills measured by minimal competency tests are learned over a number of grades, it is very difficult to hold an individual teacher accountable for the fact that at grade 8 or 10 or whatever a pupil cannot demonstrate mastery of certain basic skills. While one could quite justifiably argue that the system is at fault, the only individual who can be held directly accountable is the pupil. So long as minimal competency programs restrict the assessments of competencies to the junior/senior high school years - after two thirds of a pupil's education is over - it will be impossible to hold an individual teacher accountable for a pupil's performance on a minimal competency test.

It is not our purpose to argue here that students should not be held accountable for their learning. Rather, we are trying to point out that by default, largely as a result of the implementation procedures adopted in some states, the student assumes the primary onus of accountability. Since the potential consequence of failing a minimal competency test is grave - denial of a high school diploma - it seems important that the question of accountability be explicitly considered before a MCT program is enacted.

There are, in essence, two related questions associated with accountability. The first asks, "whom do I wish to be accountable, if anyone?". Is the concern primarily with teachers, pupils, school curricula in general, or some combination of factors? We have noted elsewhere (Madaus and Airasian, 1977) that a key aspect in the credibility of any external certification system is that the sanctions for failure (e.g., refusal of a diploma) be real. For example, one reason that drivers license exams are taken so seriously is that there is a significant first time failure rate. This failure rate gives the driver test credibility, which in turn motivates people to seriously prepare for it. If sanctions for student failure are to be applied in a minimal competency program, it is necessary that they be clearly thought through in advance of program implementation. We may ultimately decide that the student bears final responsibility for his learning, but we should decide this in advance, not let it be decided by default or by the particular way we implement a minimal competency program, as has happened in many other states.

The second question related to accountability asks, "Given the program as enacted and implemented who actually is held accountable?". The answer

to this question simply provides a check that the individual or institution one intended to be accountable is, in fact, accountable.

The question of accountability obviously is not a simple one. It involves the decision to hold some group or the "system" responsible for student learning and hence has political, economic, educational and ethical aspects. We have not argued that any particular individual, group or organization be singled out as the objects of accountability. We have suggested instead, that since a minimal competency program embodies a form of accountability, it is essential that the issues of who is to be accountable, and how accountability will be determined, be thought through before the program is instituted.

If we turn our attention now to pupils, the group most directly accountable in minimal competency programs that are linked to a high school diploma, there are a number of other issues that arise. For example, given that pupils are to be held accountable for their learning, we may ask whether a minimal competency program will have its major impact on <u>all</u> pupils or only on particular subsets of pupils. While there may be an indirect impact on all pupils, in that MCT programs may reorient curricula and instruction in schools, the major direct impact is likely to be upon a minority of pupils - those who presently are experiencing scholastic difficulties. Pupils who are already weak scholastically will find passing the MC tests most difficult. These same pupils are presently most likely to be tracked, or streamed, into remedial classes, be held back, or drop out of school. In one sense then, minimal competency programs may be seen as directed towards these scholastically disadvantaged students rather than toward the entire student population.

Given that a MCT program tied to graduation impacts most severely on pupils who already are experiencing scholastic difficulties, two further questions arise. First, "will pupils be held accountable for their performance without consideration of non-school factors which influence learning?". The answer to this question depends upon the nature of the competencies defined, e.g., are they school specific skills like the three R's or are they more general intellective applications, processes, etc. ? Clearly related to this issue is whether, given present knowledge about instruction, motivation, and remediation, we believe that teachers are capable of teaching the defined competencies to all students. If the assumption is made that all pupils can master the competencies in question, then it seems more reasonable to hold students accountable for their learning regardless of non-school factors, provided diagnosis and remediation are available when needed.

Second, given the assumption that minimal competency programs will have greatest impact on the scholastically less able student, it seems important to ask whether testing these students is most fortuitous at grade 8 or above, or whether the real problem involves their early identification. Given that MCT programs will impact most severely on a certain segment of pupils, is it necessary - or wise - to wait until junior high school to test for minimal competencies? There are other ways to identify pupils with learning difficulties early in their school careers. Moreover, a point we shall return to often, is that remediation will be more successful the earlier it is instituted. This suggests that to wait until grade 8 to remediate skills that should be mastered at an earlier grade may be inefficient and unproductive.

There are then, a number of accountability concerns which any agency contemplating a minimal competency program should think through before inaugurating a program. Among these are: "Why do I want a minimal competency program?", "Who is ultimately held accountable for student learning?", "Do the actual procedures and implementation strategies of the program place the burden of accountability on the correct parties?", "Is it not possible to diagnose, remediate, and certify minimal competency earlier than at grade eight?", "What evidence is there, or what is the basis for the belief, that a minimal competency program per se will help poor students to learn, motivate teachers to focus on the basics, change curricular emphases, or improve remediation?"

These are not easy questions to answer. However, the demonstrated need, credibility and effectiveness of a minimal competency program will depend as much upon thoughtful reflection on these questions as upon more technical issues such as the nature of tests to be used, the standards applied, the frequency of testing, and so forth. Too often questions such as the above are ignored in a frantic rush to "do something". Doing something without reflection on such issues may produce some limited immediate benefits, but it will do so at the cost of a clear conceptualization of what ends a minimal competency program is intended to accomplish, and whether a minimal competency program is the most efficacious means of accomplishing these ends. It is better to have a well thought out rationale for, and understanding of, MCT before mandating such a program, than to try to build a rationale after the fact in defense of an ongoing program.

3. What characteristics should the selected competencies possess?

The first practical problem in initiating a MCT program is the identification and selection of the competencies that will be the focus of the program. In specifying these competencies a number of sub-issues inevitably arise: which areas of the curriculum should be included; how many competencies should be included; who should identify and state the competencies; should the competencies be skills specific to the school, or should they be more associated with extra school functions; what criteria should be used in judging the acceptability of potential competencies? In this section we shall deal with these issues.

Probably the single greatest concern in selecting competencies is the answer to the question, "Is it reasonable to expect, given sufficient time and resources, that the school can foster these competencies in all children?" If students, teachers, and/or the school system are to be held accountable for performance on the selected competencies, that is, if there are sanctions associated with poor performance, it is critical that the competencies are viewed by most people as capable of attainment by all pupils. It is unfair - and probably illegal - to hold pupils, teachers or schools accountable for behaviors, skills and knowledge that are in part, or totally, beyond their direct power to control and influence. We recognize, of course, that it may be harder for some pupils to attain the competencies than others because of motivational, intellectual, physical, or emotional handicaps. Some pupils will undoubtedly require more time, more remediation, and/or more individualization of

instruction to attain the competencies than will others. But if there is doubt about whether it is reasonable to expect all pupils to attain a competency as a result of instruction - of whatever type, duration, and time - then that competency should not be part of the program.

From the point of certification, it may be insufficient to argue that "I can guarantee teaching, but not learning." A stronger case will need to be made that the competencies tested are, at least potentially, available to all students through the instructional process. Undoubtedly there will be pupils who will not master all, or even most, of the required competencies. There does, however, have to be a strong belief, or better yet evidence, that the school could foster the competencies in all pupils given the correct set of circumstances. There is an important distinction that must be kept in mind and that distinction concerns the difference between the proportion of students one expects, under optimum conditions, can successfully learn the competencies and the proportion of students who actually attain measured minimal competence. Students may fail a MC test for any number of reasons. Nonetheless the expectation must be that the student could have mastered the competencies.

The preceding argument is important because in selecting competencies schools implicitly go on record as assuming the major responsibility for fostering the selected competencies. This in turn implies that schools can foster in pupils the behaviors implied in the competencies; if they cannot, it is unfair to expect pupils to manifest the behaviors and penalize them if they don't. This argument seems simple and straightforward, but it contains an important concern.

The usual focus on the desirability of selected competencies tends to divert attention from the question of whether they are amenable to instruction in the schools. In fact, not a great deal of systematic, transferable knowledge exists about the instructional process. This is particularly true for more generalized pupil behaviors such as the ability to apply ideas to new situations, or to synthesize disparate information into a coherent whole, not to mention a wide range of social and affective behaviors.

Some have made a distinction between two kinds of pupil competencies. The first type of competencies are called "enablers" that is school specific skills such as mastery of multiplication facts, spelling, punctuation rules, etc. The second type are more generalized "competencies"; those involving the application of school specific skills and knowledge to more life-like situations such as balancing a checkbook, computing compound interest, etc. The state of the art in instruction is such that we are considerably more confident about teaching pupils "enablers", basic facts and knowledge, than we are in effecting the transfer of these skills to life situations.

While most educators would agree that basic cognitive skills and processes do fall within the domain of teachable behaviors, there appears to be less certainty that such "higher level skills" as application, analysis, synthesis and evaluation (Bloom et al, 1956) are capable of being successfully taught to all pupils. Certainly the relationship of general intelligence to mastery of these higher level behaviors belies the notion that they are readily available

to all students (Madaus, Nuttall, and Wood, 1973). Moreover, and perhaps most crucially for a competency based certification approach, it is not at all clear how pupils acquire these higher order behaviors or what types of instructional materials are most appropriate for developing them. The key issue here is not that these behaviors cannot be taught, but rather that teaching them, given the present state of knowledge, is more an art than a science. Techniques useful with some pupils are not useful with others; strategies adopted by one teacher are not readily exportable to other teachers. Certainly, teaching higher level cognitive behaviors involves a degree of complexity far removed from the simple "diagnose and prescribe" language contained in most competency-based programs. It may be that people think that the adjective "minimal" effectively excludes such higher level skills from inclusion in a MC program. However, application of basic "enabler" skills to "life like" situations could be considered a high level mental process.

If the domains of personal, social and career development are to be tested as part of a MC program, and if these domains involve more than cognitive and psychomotor competencies, then it may be particularly difficult to design appropriate instructional techniques for the development of such competencies. Even when the issue of whether the school should teach patriotism, self concept, work ethic, or job preparedness are put aside, there is a serious question about whether schools can teach such affective competencies to all pupils. Programs in these areas to date have not met with marked success, and the prospects for widespread, universally successful teaching of such objectives

remains a hope rather than a reality. There is an additional set of legal problems associated with teaching and evaluating mandatory personal and social behaviors. To date Massachusetts has avoided affective competencies and has concentrated on cognitive skills in mathematics and communications. Policy makers would be well advised to continue in this vein.

Focusing upon ends rather than means has particularly important overtones in MCT programs. In MCT programs an understanding of the limits of the state of the art in instruction and curriculum is critical, because by the explicit publication of minimal competencies, schools go on record as implying that they can, in fact, develop these competencies. In doing so schools make themselves primarily responsible for success--and failure--by implying that they can foster the competencies in all pupils. While there is no doubt that it would be beneficial for all pupils to manifest competencies which will help them cope with life in a complex technological society, there is a legitimate question of whether schools can and should assume the primary responsibility for the development of all such competencies.

In light of the foregoing, the following are recommendations regarding the selection of competencies that are to be tied to high school graduation. First, there must be reasonable certainty that the competencies are capable of being developed in all students through the instructional process. This implies that the schools have the resources, materials, and time available to give each learner an equal opportunity to master the competencies. Basic school specific skills and knowledge fit these qualifications better than do

transfer and application behaviors associated with "life skills." Second, as far as possible, the competencies selected should be specific to school instruction and independent of such factors as social class, home background, and intelligence. Obviously it will be impossible to completely disassociate minimal competence from these background factors, but it will be important to demonstrate that each competency was specific to school instruction. Third, the competencies selected should be capable of being remediated. This implies that materials, teachers, resources, etc., needed for a program of remediation are available. Finally, all of the above points assume that a curriculum for fostering the competencies either exists, or can be readily developed. It is probably better, therefore, to select competencies from existing school curricula than to start the time consuming, costly process of developing and validating new curricula. Interestingly, the minimal competencies defined in terms of "enabling skills", the three R's, are already an integral part of the elementary school curriculum but are not covered explicitly in the curriculum of the secondary school. The proposed listening and speaking competencies are not an integral part of most curricula. Life skills are presently not an integral part of the curriculum of either the elementary or secondary school. While these four recommendations do not exhaust the concerns associated with the selection of competencies for a MCT program, if they are not considered then the potential for legal challenges is great.

Two additional questions about the competencies selected remain to be discussed: "what is the appropriate number of competencies?" and "who

should select and state the competencies?" We shall consider these questions in turn. First, and briefly, there is no absolute correct number of competencies which are required in any given program. The number of competencies selected depends on a number of other factors discussed in this document: cost, teachability, availability of resources, etc. A major proliferation of the competencies occurs when MCT programs move into the area of certification of pupils on "life skills," "survival skills," or "skills needed to function effectively in society." When an MCT program attempts to certify pupils in these areas, the number of competencies that could be included in the domain of "life" or "survival" skills grows very quickly - often to an unmanageable number. The key is to start with a small number of competencies in basic, school specific areas such as reading, writing and mathematics. By focusing on these basic areas, as they are taught in schools, the number of competencies can generally be kept to a reasonable, manageable number.

We shall have more to say on the issue of who should state the minimal competencies later when we compare three specific models for implementing a minimal competency program. Suffice it to say at this point that an important characteristic of any set of competencies is that it be politically viable. In essence, the competencies need to be accepted by various publics at the state and/or local level. The primary way to attain political viability is to involve representatives of various publics in the selection process. For the competencies to be seen as valuable, and for the program to be viewed as viable, parents, teachers, administrators, employers, etc., from the community must feel that the program is, in some sense "theirs." Further, the Florida

experience with MCT clearly shows that the program will adversely impact on the poor, on certain minority groups and on pupils whose first language is not English. Individuals representing these groups therefore must be involved in the selection process. Court challenges to a MCT program could turn on a lack of minority group participation in the determination of the competencies to be tested as part of the high school certification process.

There are many ways in which community involvement in the process can be carried out, including community needs assessments, large group meetings, smaller groups identifying competencies from which more representative groups select, and so forth. However it is done, broad based community participation in selecting (or stating) competencies is absolutely necessary.

Teacher involvement should be the keystone of any MCT program. Teachers should be involved in the specification of the competencies and in the preparation of teaching materials to obtain these competencies. Teacher involvement not only insures the good will of those who must guide students in attaining the competencies, but also helps to insure that the competencies are relevant, the teaching materials appropriate, and that the examinations are reasonable and valid.

4. When should pupils be tested for minimal competence?

One of the parameters set by the Policy Subcommittee for our analysis was that we take grade 8 as the grade at which testing would begin for the

certification of minimal competencies. This does not, however, preclude the possibility of testing particular competencies at earlier grades. As a general rule of thumb, testing for minimal competency should occur early and often. Some form of testing should occur early to identify learners at risk; those who are likely to require some form of remediation to attain the competencies. The earlier such learners are identified the more likely it is that remediation will be successful - if only because there will be more time for remediation. The actual certification tests should occur often because there are individual differences in the rate at which pupils learn; all pupils are not equally ready to demonstrate competency at the same point in time. Further any minimal competency test will be an imperfect measuring device and this, coupled with the fact that some students are poor test takers, means that some students will be misclassified by the test. For these reasons multiple opportunities to sit for the certification test must be made available. Further, the high school diploma is tied to passing the minimal competency tests - and since the high school diploma is a valuable commodity in society - pupils have to be given a number of opportunities to demonstrate competency.

There is no specific rule for the frequency of testing, but many states and districts test at least once a year from grade 8 onwards. Others test twice a year over the same period. The frequency of testing will in part be determined by a number of practical considerations including administrative, scoring and test construction costs and the time required for testing. If

performance, rather than multiple choice paper and pencil tests are used, cost and time factors may prohibit testing more than once a year.

A potential negative consequence of testing "early and often" should be pointed out. That is because the tests are labeled minimal, if they are administered early, it is possible that those who fail the first time - or those who repeatedly fail on subsequent administrations - will be tracked or streamed within the school in a way that stigmatizes them publicly. Will these students be cut off from their peers? Will they eventually become school leavers? There is a danger then, that by identifying pupils at risk, or pupils who fail the minimal competency test, that a self-fulfilling prophecy of failure is set in motion.

The Florida experience with MC testing has shown that poor black pupils are a population at risk. In communities with a large minority population this fact poses another problem associated with identifying and treating minority group children at risk in order to help them attain the minimal competencies required for graduation. That is, once these children are identified through testing, placing them in special classes for remediation, or pulling them out of regular classes for help, could be challenged in the court under Title VII of the Civil Rights Act on the basis that the program is re-segregating the school. When the courts have been confronted with data showing that test results are used to form groups for differential treatment along racial, ethnic, sex and economic lines then they generally bar, or severely limit, this practice (Tractenberg & Jacoby, 1977). Thus if test(s) used to identify those at risk

have a disproportionate racial impact the courts could rule that the testing practice be discontinued.

Schools and teachers will need to guard against making some pupils "second class students" because they have been identified as being at risk in terms of passing the minimal competency test. Obviously there is a dilemma here. We need early testing in order to identify and help those for whom the prognosis of passing the MCT is not good. On the other hand, schools will need to be extremely careful that this early identification does not lead to pejorative labeling, grouping or pull out procedures, that violate the principle of mainstreaming which is presently operative in the Commonwealth. There are serious ethical, practical, and legal implications embedded in attempts to diagnose those at risk; state and school officials need to be cognizant of these problems in developing their remedial efforts.

Another problem related to the timing and frequency of testing is whether, once having demonstrated competence, say in grade 9, the pupil at grade 12 must once again demonstrate the mandated competencies. Pupils change, forget, lose interest, etc., over time and there is no guarantee that passing a minimal competency test at grade 9 can be taken as evidence that the competencies are still operative at grade 12. Most states and school districts certify pupils when they pass the MC test(s) for the first time, regardless of grade level and do not demand that the competencies be demonstrated again at a later date.

5. How should special pupil populations be dealt with?

There are three pupil populations at risk in MCT programs: those who are handicapped, those whose native tongue is other than English, and those who are poor and from minority groups. The problems associated with testing these populations for minimal competence are primarily legal, in the form of discrimination suits associated with the relevance, items, language, or mode of testing associated with minimal competency tests.

As far as handicapped pupils are concerned, it needs to be recognized that there are many types of handicaps which may for convenience be grouped into the three general categories: physical (blindness, deafness, etc.), emotional, and mental. In considering the efficacy of applying minimal competency tests and standards to handicapped pupils, it seems clear that there are potential problems no matter what strategy is adopted. If such pupils are exempted from minimal competency testing and certification - and hence receive a diploma that is different from the diploma awarded to those who pass the minimal competency tests - they may have a legitimate claim of discrimination on the basis of being handicapped. If, on the other hand, handicapped pupils are required to demonstrate minimal competence, there likely will be legal challenges to the testing procedures and standards.

There is no easy answer to the question of how to deal with handicapped pupils in minimal competency programs. Public Law 94-142 requires individualized programs for handicapped pupils. In the spirit of this law, it seems most reasonable to treat each handicapped child on an individual

basis, rather than attempting to implement a blanket policy toward the handicapped. It does seem reasonable, however, to consider the physically handicapped as presenting a somewhat different problem than the emotionally or mentally handicapped. It may be possible to hold the physically handicapped - depending on the nature and severity of the handicap - to the same standards as their non-handicapped counterparts, but to utilize different testing procedures to overcome the particular handicaps of any given learner.

For the emotionally and mentally handicapped the problem is more difficult. There are many degrees of emotional and mental handicap and it seems unreasonable to lump all degrees together to define a general policy for such pupils. Individual needs, options, and decisions may - and probably should - be made for this population. Certainly this is an area in which additional counsel should be sought before a decision is made.

The question of foreign-speaking students is similar in many regards to that of special needs pupils. If proficiency in the English language is a prerequisite for a diploma, it is necessary to be able to demonstrate that foreign speakers have had sufficient opportunities to learn the competencies in English. Obviously this requirement poses a long range problem, since English speaking pupils have had all of their instruction in their native tongue. Foreign speakers have had some form of bilingual instruction which for many make tasks involving English more difficult. Of course the option of a foreign language test is always possible, but if facility in the English language is perceived to be a necessary competency for high school graduation, foreign-language tests are

not viable. It appears that the problem of foreign speakers vis a vis minimal competency testing will never be solved in any general way since it requires certain types of educational experiences which permit valid arguments that foreign speakers have had sufficient school-specific opportunities to pass a test in the English language. Again, more in depth counsel should be sought on this question.

As we noted in the previous section, the Florida experience with MC test(s) linked to graduation clearly shows that poor black pupils are a population at risk. The issue of disproportionate racial impact also raises issues of test bias and test fairness. If there is a disproportionate racial impact the charge that the test is biased or unfair must be anticipated. Item analyses could be used to discover questions which prove unusually difficult for minority students. However the technique may not be able to eliminate all items which differentiate between races. Another "bias" charge might be that the tests are in standard English and poor, minority students who are familiar with a different idiom are at a disadvantage (Turnbull, 1977). This issue has to be met head on. Tests of basic skills measure the ability to read, compute and apply standard English to solve problems (Turnbull, 1977). We need to know if poor children, black and white, cannot handle basic English. If they cannot we need to have this type of information to change this deplorable situation (Learner, 1977).

However to attack as discriminatory the MC test(s) because it is in standard English smacks of "slaying the messenger." It is the instruction that is at fault

and not necessarily the test(s) per se. Further, to suggest that MC test(s) should be designed to measure survival skills needed in the ghetto also misses the point. Because the test(s) show poor blacks have trouble with skills thought to be important in the predominant culture does not mean the tests are necessarily unfair or that they need to include skills needed in minority cultures.

The issue of test bias is a complicated, technical, emotional and political issue. It needs to be addressed before the MC test(s) are built and administered. It needs to be addressed by directly involving representatives of poor minority groups in the entire process of developing the MCT program.

In general, it seems most reasonable to handle special student populations on a case by case basis until some parameters for a more general policy can be worked out and justified both educationally and legally. This strategy implies giving students options, consulting parents, making special provisions for testing, and the like. In many respects, the issue of special needs pupils is the one issue addressed in this report on which it is most difficult to make generalizations about procedures and viable options.

6. How much need will there be for remedial programs?

The first question regarding remediation that must be answered in considering a minimal competency graduation requirement is whether or not remediation will be a required component of the minimal competency program. We feel that the answer to this question must be an unequivocal "yes." One

cannot require all students to attain minimal competencies in prescribed areas without assurances that every effort has been made to enable all students to learn these competencies. Thus, any student who fails the minimal competency test, or is identified as being at risk of failing the test, must be provided with remedial help in an attempt to bring him/her to a satisfactory level of competence.

Given that remediation is required, the next issue that must be addressed is, "How will remediation be carried out?" Closely intertwined with this question is the further question, "When will remediation be carried out?" With regard to the latter, waiting until a student takes and fails a minimal competency test for high school graduation at grade 8 seems to be too late for a number of reasons. First, if one waits until grade 8, the remediation efforts become the task of the upper junior high or senior high school teachers. Junior and senior high schools presently are not structured to handle large scale remediation. Most secondary teachers are certified to teach a particular subject and are not trained to provide remedial instruction. A number of secondary teachers might therefore have to be retrained or elementary staff would have to be shifted so that teachers capable of providing remediation were available at the secondary level.

A second reason that grade 8 may be too late to begin remediation of minimal competencies is that the problems associated with a student's failure to master the minimal skills may be so ingrained by then that remediation would be very difficult. Early identification of the minimal competencies

needing remediation is necessary, and generally the earlier the better. To undo the cumulative failure of six years of elementary school will be very difficult and expensive at the high school level.

Third, even with early identification, programs might need to be developed to meet the needs of pupils at risk. Title I children probably constitute one segment of the population at risk in terms of meeting MCT requirements. Most pupils in Title I are poor and well behind their peers in achievement. Unfortunately Title I evaluations do not lead to a sanguine prognosis that the achievement level of the majority of children served by Title I will be dramatically altered. Nonetheless the state might try to explicitly re-direct Title I programs to serve children at risk in terms of minimal basic scholastic competencies.

Finally, there is an ethical imperative to provide remedial instruction as early as possible. We have argued throughout that some assurance is necessary that students taking the minimal competency test have a reasonable chance of passing the test. Unless attempts are made to identify students at risk in terms of minimal competencies prior to their taking the test, and then provide them with remedial instruction in the areas where they are weak, it is not ethical to require them to pass a minimal competency test in order to receive a diploma.

To identify children at risk may involve more in the way of diagnostic testing than is presently done in schools, and perhaps there will be a demand for more core evaluations and prescriptive programs similar to those done

under Chapter 766. In fact a MCT program will probably result in an increase in requests for "core" evaluations and prescriptive programs under the provision of Chapter 766. If this happened it could substantially increase the costs to local districts for Chapter 766.

We must be careful to keep minimal competency testing in perspective with regard to the total educational process. Remediation in basic skill areas is not going, all of a sudden, to become an issue because students have to pass minimal competency tests. Remediation in the basic skill areas is already an integral part of the educational process in the early elementary grades. In those districts where remedial services are provided to only a small number of students, especially at the junior/senior high school level, there is the risk that the district will have to provide remedial services to a greater number of students as a result of a minimal competency testing program. In those districts where remediation begins in the early grades and is continued for as long as is necessary, minimal competency testing should not significantly increase the numbers of students requiring remedial services at the secondary level.

The question, "Will a significantly greater number of students be identified as in need of remediation as a result of a MCT program than are presently identified as in need of remedial services?" must be examined further. If the answer is yes, there are cost and staffing problems to consider. The answer to this question is tied in part to the level of competency that is required of students. Clearly, if half of all students fail the minimal competency

test, more students would need remedial instruction than are currently being provided such services. If half the students fail, however, the requirements for passing would be suspect. If "reasonable" standards are adopted, the answer to the question posed here is probably "no, the numbers will not increase substantially." Florida provides an indication of the impact of a MCT program on present remedial services. In that state 36% of the pupils failed the competency test in arithmetic, 8% failed the language test. As a result Florida earmarked 10-12 million dollars for remediation in arithmetic. As we will discuss subsequently, this negative finding may simply be an artifact of the item difficulty of the minimal competency test used to measure arithmetic skills. Nonetheless, in the state of Florida minimal competency test results were interpreted (we think incorrectly) as indicating that large sums of money were needed to institute a program of remediation.

The institution of a MCT program may provide more focus to the remediation efforts in the schools. Such a program necessitates a more formal approach to remediation than may already exist in many schools. As we argued above, remediation in the specific skill areas that are to be tested is a must. Documentation that these areas were taught, and that the student either attained the required competencies, or was given every opportunity to reach the desired level of attainment, would also be necessary to answer possible legal challenges. Procedures for monitoring student achievement in the required skill areas, or their prerequisites, beginning in the early grades, might very well have to be adopted. These procedures - probably a combination of teacher designed,

district designed, and commercially available tests - could be incorporated into the process of identifying students needing remediation. Thus, while a MCT program need not lead to a massive increase in the remedial services already offered by a school system it may lead to a reorganization of those services in terms of their focus, record keeping and testing procedures in order to document that remediation has in fact been provided to any student in need.

Assumptions made when considering the remediation issue should be made explicit. The first assumption is that all students who need remediation can be identified. The second assumption is that the specific areas in which these students need remediation can also be identified. It is not enough to be able to say that a student is having difficulty in reading, the specific problem(s) underlying that difficulty have to be identified if specific instruction is to be provided. If the state moves to certify "life skills" identification of the specific factors associated with failure to master these skills could be a problem.

A third assumption is that we know how to remediate the problems that have been identified. This assumption is not always tenable. There are students currently receiving remedial services who benefit and some who do not. The question of whether it is enough merely to provide remedial services to all students in need, or does the state wish to further require that all students receiving remedial services be brought to a level of functioning where they can pass a minimal competency test, is one that educators and policy makers need to address.

Finally, what is to happen to those students who have reached the end of grade 12, have been provided with remedial services, and still cannot pass a minimal competency test? Should the school district assume a further responsibility for this student and offer him/her the opportunity to stay on in school in an attempt to pass the test at some future time? Should a certificate of attendance be granted to this student in lieu of a high school diploma? If so, what does such a certificate mean in terms of employment and life chances? Will the frustration level of students who repeatedly fail the minimal competency test in spite of attempts at remediation become so high that they will leave school and thus increase the number of dropouts? All these questions are worthy of this committee's further consideration.

In summary, the issues of when and how to remediate in the context of a minimal competency testing program appear to be very similar to those remediation issues that currently exist in our educational systems. The answer to the question of when to remediate seems to be "as early as possible." Early identification of students needing remediation in the minimal competencies is crucial. The answer to "how to remediate" is not quite as simple. It is not clear that we know how to remediate all students in all situations. It seems that minimal competency testing will impact on the "how" aspect of remediation by linking remediation directly to the required minimal competencies. Further, because of legal challenges it will be necessary to carefully document not only that remedial services were provided to students in need but also the specifics of the services received.

7. Will Minimal Competency Tests impact on the curriculum?

What will be the effects of implementing a MCT program linked to high school graduation on the school curriculum? The answer to this question is a function of three factors:

1. the perceived importance of the competency test
2. the extent to which the minimal competency test measures skills different from those taught in the existing curriculum, and
3. the amount of lead time or phase-in time between announcing the requirement of a competency based exam and the date on which that requirement takes effect.

With regard to the first factor, the minimal competency tests will have little or no impact on the existing curriculum if the tests are perceived as unimportant, or lacking in significance. On the other hand, if it is very clear that the diploma depends on passing the tests, and further that not everyone passes the test(s) the first time, then de facto the test(s) will become important and be taken seriously. Passing the test(s) becomes crucial to students who value a high school diploma. Further assuring that most students attain the minimal competencies becomes important to school board members, superintendents, principals, and teachers, since the perceived value of the education the district provides now rests in the balance. Parents and employers who value a high school diploma will also view the competency test as important. Thus, the importance, or perceived importance of the MC test(s) is a necessary first ingredient in its impacting on the curriculum.

The second factor that will influence the impact that minimal competency test(s) have on the curriculum is the extent to which the test(s) measures skills different from those fostered by the existing curriculum. A curriculum which emphasizes most, or all, of the competencies measured by the minimal competency exam would not be expected to change greatly as a result of the test(s) - unless there was a high failure rate. Thus, test(s) covering speaking and listening competencies, or tests covering "life skills," could be expected to have an impact on the curriculum, since these areas are not presently covered directly and/or in depth by most curricula. Further, if the curriculum was not changed in these cases, it would be impossible to justify administering the test to students. The test would simply not be a valid measure of what the students had learned in school. The test would not be fair to students, and would be open to legal challenge. Prior to adopting a minimal competency program, a survey should be conducted to determine how curricula differ from district to district, and how the proposed competencies differ from the areas already emphasized in various curricula throughout the state. Some determination of the impact of a minimal competency program on various curricula could be made from the results of such a survey.

Of course an assumption is that the competencies, or objectives on which the minimal competency test(s) were based, were clearly spelled out for administrators, teachers, students, and parents so that competencies not emphasized in the existing curriculum could be readily identified. This implies that the competencies are published and that the test(s) are available to the public after each administration. (We shall return to this latter point in the

section on test security.) Jencks (1978) points out that until the public sees the exams there is no way for it to know how well they measure the things it thinks are important and should be certified before a diploma is issued. He goes on to point out that teachers who do see present commercial achievement tests argue that the tests don't measure things they want students to do, or to know about. Some in the testing profession cavalierly dismiss this reaction as teacher defensiveness. However, until tests are available for public scrutiny such a charge of defensiveness cannot be intelligently countered.

Also note here that a legal issue of concern may be the injury (i.e., denial of a diploma) to students because of factors largely beyond students' control. Large numbers of students might be denied a high school diploma, despite learning what their school emphasized, simply because the emphasis of the minimal competency test differed from the school's curricular emphasis. Students would, in essence, be held accountable when one might quite reasonably argue that the real burden of accountability should fall upon educators who determine the school's curricular orientation.

Curriculum change takes time. This brings us to our third issue, the lead or phase-in time between requiring that a student pass a MC test(s) to graduate and the date at which the requirement takes effect. To mandate in June that minimal competency testing will begin the following September and expect the required competencies to have an impact on curriculum is unreasonable. On short notice (conservatively estimated to be less than 2 years), the only test that could validly be administered would be one which was built to measure the skills being taught in the existing curriculum.

If one objective of a MC program is to effect change in the curriculum, much more time would be needed for a number of reasons. First, the competencies on which the tests are to be based would have to be disseminated to school staffs. Second, enough time would have to be allowed for the necessary changes in the curriculum to take place. Again, the necessary changes could vary from none, in the case where the curriculum already emphasizes all, or most, of the areas specified in the list of competencies, to a great deal of modification in the case where the existing curriculum emphasizes few of the required competencies. In the ideal case, only students who began school, i.e., entered first grade, after the modified curriculum was in place, would be required to pass a minimal competency test in order to receive a high school diploma, since only these students would receive the revised curriculum through their entire school experience. Such an idealized situation requires a lead time on the order of 11 to 12 years - 2 to 3 years for curriculum change plus 8 to 9 years (assuming testing in grade 8 or 9) before the test can be administered for the first time. This amount of lead time does not seem reasonable, but the issue of how much lead time is necessary is still problematical. However, where necessary, sufficient time must be allowed for both curriculum change, and for that change to impact on student learning.

Thus far in this section, we have dealt with the prerequisites for a MCT program to have an impact on curriculum. We shall now speculate about the types of change that may result. One possible change could be an increased emphasis on the skill areas dealt with in the competency tests. Given such

an emphasis and further given a fixed amount of time, a corresponding decrease in emphasis in other areas of the curriculum, e.g., foreign languages, physical education, the sciences, art, music, etc. might result. The questions that must be asked here are, "Will a minimal competency testing program narrow the scope of the school's curriculum so that the competency areas are the primary focus and other areas are peripheral?" and "Do we want this to happen?".

Another possible impact of minimal competency testing is that it could lead to a homogenizing of curriculum. All students may be required to concentrate on the competency areas, using the same curriculum, until a determination is made that they have the necessary skills to pass the minimal competency test. Once that determination has been made, the student would then be allowed to branch out into other academic areas. The implications of such a curricular change are worthy of note. Questions arise such as, "What will happen to non-traditional educational settings (e.g., ungraded or open classrooms)?", "Are students differentially affected depending on their academic ability?"

It seems that for the majority of students, a narrowing or homogenizing of the curriculum until they attain the minimal competencies would not severely affect their curriculum. These are students who are bright enough and motivated enough to master the minimal skills fairly easily. There is another small segment of the population who, for reasons of limited ability, emotional problems, etc., may never master the minimal competencies. The remaining small segment of the school population is comprised of those students who we may categorize as borderline. Given the proper instruction, enough time and

encouragement, many of these borderline students may be able to pass a minimal competency test. The major question that must be asked then is, "Are all the efforts that are expended in developing competencies, and tests of these competencies, revising curricula, and providing remedial services, really directed to only this small segment of the school population who are judged to be in the borderline or questionable group?" Related to this question is, "Are these the students we want to focus on in a minimal competency testing program?" Pilot testing of the proposed instruments, together with input from teachers, could provide an indication of the number of students who fall into this borderline group.

In summary, the prerequisites necessary for a minimal competency testing program to impact on curriculum are that the test be perceived as important, the competencies required differ from those currently being taught, and the lead time be sufficient for curricular change and subsequent student learning to take place. Potential changes in curriculum could be increased emphasis in the required minimal competency areas and decreased emphasis in other areas, homogenizing of curricula so that all students concentrate on the competencies in the same way, and differential impact of curricular change on students depending on their academic ability. Unfortunately, the impact of minimal competency testing on school curricula will not be completely evident until the program is well underway.

8. How are standards for passing a minimal competency test to be established?

"How does one set standards or criteria for determining whether or not a student has passed a minimal competency test?" This issue is complex and raises a series of questions related to the setting of standards which will be discussed in this section.

One question that arises almost immediately is, "Should the minimal competency test be norm-referenced or criterion-referenced?" A norm-referenced test is one in which the scores are interpreted in terms of the pupil's performance relative to that of a norm or comparison group. Norm-referenced scores generally take the form of grade equivalents, percentiles, or standard scores. Some states and districts have decided to use a commercially available, standardized, norm-referenced test as their test of minimal competency. These states have generally set the average performance of students at a particular grade level as the standard they are willing to accept as indicative that the student has mastered the minimal competencies. For example, a student might have to perform at the level equivalent to the average performance of students beginning the ninth grade. In other words, a grade equivalent score of at least 9.0 becomes the score a pupil must obtain to be certified.

There are a number of problems associated with this practice. First, the commercially available norm-referenced test de facto defines and selects the competencies being certified. It may be extremely difficult to find an available test that covers the skills which the school wishes to certify.

Further, standardized, commercial norm-referenced tests of basic skills are generally not designed to be administered to high school students but instead are meant for use at the elementary grades. In such cases interpretations of the norm tables, that is using grade equivalent scores from a test designed for elementary rather than secondary students, will yield misleading results. Grade equivalents for elementary school tests that extend into the high school years, are the result of extrapolations based on rather dubious assumptions and have little or no meaning. An eleventh grade student who obtains a grade equivalent score of 9.0 on a basic skills test designed for elementary pupils cannot be said to be functioning at a level comparable to the average student at the start of ninth grade. Unfortunately this is exactly the meaning that would be implied if a 9.0 grade equivalent were used as the standard for certification.

It would be possible to design and norm a test of basic skills for high school pupils. In such a case the grade equivalent score of 9.0 would be referenced to the average performance of ninth graders. However, the question would inevitably arise as to why set the passing score at the 9.0 level, why not the 12.0 level? That is, why accept the average performance of a ninth grader as the criterion for a high school diploma, why not the performance of the average 12th grader? There is a very good technical reason for this. A grade equivalent of 12.0 would be the median score for all 12th graders in the norming population. This means that one-half the 12th graders fall

above this point and one-half fall below. Thus to set the cut score at 12.0 would mean that one-half the 12th graders and even fewer juniors and sophmores would be certified. Thus a grade equivalent of 9.0 is politically more acceptable. To circumvent the problem of using a grade equivalent score as the criterion for passing a MCT, it may be wise to avoid the language of grade equivalents altogether.

Now let us turn to criterion-referenced tests. These tests do not reference a student's performance to that of a norm group. There are two meanings associated with criterion-referencing. The first is that a standard of performance is predetermined, and students meeting or exceeding that standard pass, or are classified as having mastered the competencies, whereas students who score below the predetermined cut-off score fail. The intended purpose of a criterion-referenced minimal competency test(s) is to separate students into two groups, those who have mastered the competencies and those who have not. This type of criterion referencing is not without its problems, however. The major problem being how does one determine in advance what an appropriate standard or cut score should be. Witness the Florida experience in which 36% of the juniors in the state failed the math section of the competency test and 8% failed the reading and grammar section. Either the cut scores on the two tests were not comparable or students' skill in math are much poorer than they are in English. It is unclear which is the case. Nevertheless, Florida has decided to place the bulk of their remediation efforts into mathematics,

a decision which is seemingly warranted based on the relative percentage of failures in the two subjects. We shall address this point presently, however.

There is another problem in interpreting criterion cut scores assigned to a minimal competency test. Some people think that a score of 80% on a minimal competency test gives us information about what a pupil can do in some absolute sense. This assumption is not completely correct. It is true that the percent of items answered correctly by a pupil is independent of how his/her classmates do on the test and hence is not norm referenced. However, unless all the items on the test are clearly linked to a very well defined skill domain, the percentage score tells us nothing about what a student can or cannot do relative to a given domain in any absolute sense. A score of 80% can be said to be criterion referenced only in the sense that some arbitrary criterion or cut score has been set. It is not criterion referenced in the second widely accepted meaning of the term, that is, that the test is designed and constructed in a manner that defines explicit rules linking patterns of test performance to behavioral referents (Jackson, 1970). This point needs to be kept in mind when publicity about minimal competency tests start to be formulated by the sponsoring agency. It could be misleading to bill the competency tests as criterion referenced unless there is a link to a clearly specified skill domain.

To compound further the problems associated with establishing standards, the question must be asked, "Do we want a single total test score indicating competence, a score in each major subject area (math, reading, writing,

etc.), or a score on each individual competency?" A single total test score, either across subject areas or within a single subject area, may be suitable for making a pass/fail decision, but is virtually useless for the purposes of diagnosis and remediation. The total test score approach allows for the possibility of a student correctly answering all items related to certain competencies, and incorrectly answering all items related to other competencies. When this happens should that student be certified as having attained the required competencies? Further if one test covers heterogeneous competencies and a single cut score is used to determine pass/fail then questions about the meaning attached to that score can be raised. In such a situation the agency would not be certifying each competency but instead that the pupil answered a certain percentage of heterogeneous items correctly. Such a procedure would hide more than it shows. To take a simple example, a minimal competency test of 80 items covers four competencies which we shall call A, B, C, and D with each competency measured by 20 items. If the cut score is set at 80%, a not uncommon figure - then a pupil need only answer 64 items correctly to be certified. One student could answer all the items associated with competencies A, B, and C, and only 4 associated with D. Another could answer all items associated with B, C, and D and only 4 associated with A. Both pupils would receive the same overall cut score of 80% and be certified as having demonstrated minimal competency. Obviously both students are deficient in separate areas.

If the agency wished to certify each competency identified by panels as desirable, then a separate test for each competency would have to be built. This would be fiscally expensive and could quickly become prohibitive in terms of the testing time required of candidates. A competency by competency scoring procedure also increases the record keeping task.

An alternative would be to build a separate test for each major class of competencies (i.e., computation, reading, writing, listening, etc.). The drawback to this approach is that a pupil could pass 2 or 3 tests and fail the remainder. The certifying agency might have to certify each test but require that all tests be passed before a diploma was issued, thus complicating record keeping.

Separate scores for separate competencies would give more information in terms of performance. However it should be pointed out that even if separate scores are used, their diagnostic value in terms of planning remediation programs for an individual or group is slight. A separate diagnostic test would have to be built to pinpoint specific weaknesses. The construction of diagnostic tests involves developmental validation costs over and above those associated with the construction of the certifying exam.

There are no clear answers to the issue of one total score versus scores for each competency. The better approach should stem from the purpose of the minimal competency testing program. If the purpose is simply to categorize each student into either the pass or fail group without regard to diagnosis and remediation for those who fail, and without regard to whether specific compe-

tencies were mastered, the total score approach may be defensible. If the purpose is to do more than that, the competency by competency approach is more defensible. As stated, setting the cut-score (or standard) is problematical with both approaches and no empirically validated solutions exist.

The major point here, however, is that setting a cut-score, be it norm or criterion referenced in nature, is largely a subjective, judgmental process. Much has been written about statistical approaches to determining the particular score that will minimize wrong decisions, i.e., categorizing a student as mastering the competencies when in fact he/she has not or vice versa. Such approaches presuppose that someone knows what the correct categorization of the student is in the first place and obviously, this is not the case or there would be no need for the test. Although the approaches are technically sound, they ultimately rely on the ability to make subjective judgments regarding the relative loss incurred in erring in one direction as opposed to the other. Comparisons of alternative statistical approaches, all technically sound, yield very different results. Thus, although at first glance it appears as if the statistical technology is available for determining cut-scores, there is no consensus on any one approach and all approaches rely on subjective judgments of some sort.

Even if the statistical approach to determining cut-scores were resolved, other problems still exist. What if a cut-score which minimized the number of wrong decisions results in 70% of all students failing the test? The cut-score

may be sound on technical grounds but is it politically, economically, and educationally sound? In reality, part of what enters into the determination of a cut-score is the answer to the question "What percentage of students can we reasonably fail?" There are obvious costs involved in failing a high percentage of students including monetary, political, and of course educational costs. Increased remediation services would result. Increased frustration for the student would also result. There are issues related to failing too few students also, many of which have been discussed elsewhere. These issues include relegating the test to meaningless because almost all pupils pass. This result may, in turn, produce criticism for having set standards too low.

A cut-score could be determined by administering a representative test to a representative sample of students to see which scores correspond to various failure rates. Once an acceptable rate is decided on, that score could become the standard. There are problems with this approach, however. Those who dislike norm referencing would argue that the approach is ludicrous, that the standard should be set independently of how some "norm" group of students performs. However, to set the standard arbitrarily can result in a massive failure rate. Suffice it to say that the task of setting standards is very complex and to a large degree subjective.

A final issue related to setting standards is that of differential failure rates for different ethnic, racial, sex, and location (i.e., urban vs. suburban) groups. If a common cut-off score is used for all students, there are sure to be different percentages of students failing when the results are examined

by some of the above strata. This differential failure rate will probably become a legal issue. The ramifications may be even more widespread. Issues such as equality of educational opportunity, test bias, and urban/suburban school comparisons are highly probable. A question that should be addressed, therefore, is "Should there be one standard for all students or differing standards for different types of students?" If different standards are adopted then logical questions about the implications of having different standards for different groups arise.

In summary, the issues surrounding setting cut-scores are very complex. Decisions about whether to use norm- or criterion-referenced tests have to be made. The manner in which cut-scores are determined for the test(s) can vary. Setting criteria for passing is a highly subjective, judgmental process. There is little agreement on an appropriate statistical technique, and the issue is broader than the unavailability of an agreed upon technology. Political, cost, and educational issues all come into play when deciding on standards. There is also the issue of whether a single total score in a skill area is sufficient or whether scores for each competency within a skill area are necessary. Finally, the issue of differential failure rates by different ethnic, racial, sex or geographical groups when a single criterion is used for all students needs to be considered. The problems associated with setting standards are difficult ones. The solution to many of these problems appear, at present, to rest more with ad hoc decisions than with existing, validated technical procedures. In the final analysis, setting standards is

highly subjective and risky, and only after considering all the problems and implications inherent in setting standards can reasonable standards be established.

9. What are the problems of test security in minimal competency testing programs?

One problem that must be anticipated in setting up a minimal competency testing program is that of test security. It is possible to set up a completely secure testing program in which the same test or large portions of it can be used every year. While many college and graduate admissions tests are kept secure from year to year, in minimal competency programs it may not be possible or desirable to build completely secure tests.

The situation of a test linked to a high school diploma is not comparable to those programs run by professional testing and certification agencies. In the latter case, the examinee voluntarily elects to sit for the exam. The examining body provides a service to the examinee for a fee, part of which goes to defray the expenses involved in building a test, running the program, and reporting results to the examinee or to schools designated by the examinee. In the case of certifying boards in medicine, nursing, and other professions, the applicant is applying for membership in a private organization. The applicant pays a fee to sit for the organization's certification exams which once again goes in part toward test development, maintenance and program costs.

The MCT program is quite another matter entirely. First, secondary pupils do not elect, but for all intents and purposes are required, to sit for the test(s). Second, the test(s) are paid for out of tax funds. If a pupil fails the exam, and is denied a diploma until he/she can pass it, it is quite conceivable that the applicant will want to question parts of the exam. Under the freedom of information laws the student might be able to demand to see the test. Further as we noted previously the political viability of the test(s) may depend on teachers and parents inspecting them to see if they are measuring things of value.

External testing programs tied to high school certification historically have not been totally secure. In past years, old New York State Regents exams could be purchased in bookstores throughout the state. In the British system, each year the exams move into the public domain once they have been administered. In fact the tradition of past exams in external testing programs dictates how pupils study, what they study and how teachers teach.

While a MCT should not strive for completely secure administration, nonetheless the agency must make sure that the test is secure before it is administered so that a fiasco such as occurred recently with the fire fighters exam does not occur.

Attempts to cut down on testing costs by having a completely secure test(s) to be used year after year therefore is probably not possible or desirable. Given this fact, test development will be an ongoing yearly affair, not a one shot expense. If the state were to give the LEA's a list of tests to

choose from, that list might have to include a new test(s) or form(s) of the test yearly. This expense should not be dismissed lightly. Nor should the state listen to arguments from commercial publishers that they will be able to keep a test secure over the years. Once the taxpayer's monies are involved, it is only a matter of time before someone asks for the test's release and the state will probably have to accede to such a request.

Even if the certifying test is made public each year after it has been used, it is still possible to try out items for the next year's examination when administering the exam. Questions are sampled across pupils so that several hundred questions could be tried out for possible use the following year. These questions are not scored as part of the certifying exam however. Thus when the test is released to the public only the questions used to arrive at the certification score need be published. This system of sampling the next year's items as part of the certifying exam also permits the test constructors to equate exams from one year to another. However each year the standard set for passing may need to be reviewed.

Another issue under test security is that of cheating. Techniques for the detection of cheating will have to be considered. Techniques to assure that the person sitting for the test is in fact the person to be certified would have to be developed. This would best be handled by the LEA's. It should be noted that if local teachers are involved in administration it would be next to impossible to keep a test completely secure over the long term. This does not mean that groups of teachers can not be involved yearly in test construction.

If a person fails an exam the temptation to have a ringer sit for the exam a second time will be great. It happens in programs run by ETS and CEEB. There are statistical ways to detect possible cheating of this type and these need to be explored. Cheating in the exam in the form of copying will have to be handled by proctors. Again, given the possibility of legal action, proctors may not wish to challenge pupils but to mark where the suspect is seated and then a statistical analysis of his/her response patterns with those of neighboring candidates' can be made. Legal advice also needs to be solicited if the state wishes to void an individual's test for cheating. Mechanisms for review would have to be set up.

10. What are the issues of test validity?

A central issue in any MCT program is the validity of the test(s) used to certify attainment of the competencies. Already there is litigation involving validity of the Florida MC tests. There is every reason to expect legal challenges from those denied a diploma based on test performance. Recent court decisions indicate that when a challenge to the use of tests arises the courts rely on the 1974 Joint APA, AERA, NCME Standards for Educational and Psychological Tests in deciding whether the test, and its use are acceptable. Therefore any MCT's used in Massachusetts need to be evaluated in terms of the 1974 Standards. Any agency charged with building and administering MCT's must be sure that such tests meet the 1974 Standards. While we shall not go into a detailed technical discussion of validity, certain issues concerning the validity of MCT will be discussed. If the MCT's are used to certify that the pupil has acquired certain

basic scholastic skills in areas like computation, reading, writing, speaking, listening, etc., then the test(s) must be shown to be content valid. To demonstrate content validity it is first necessary to define or specify the univerise or domain of interest (i.e., computation), then describe or specify the skills and/or behaviors that are part of the domain (i.e., adding two digit numbers that involve carrying, etc.). Once the domain has been specified one must then offer evidence that the items included on the MCT's are relevant, representative samples of the behaviors and skills included in the domain. Thus, if one were to construct a test to measure the student's mastery of basic computation skills, the processes (addition, subtraction, multiplication, and division) and behaviors (i.e., adding two digit numbers with carrying) to be included in the domain would have to be clearly specified and delineated. Domain definition is perhaps easiest in the area of arithmetic. In the area of reading comprehension, the difficulty of the material to be understood would have to be specified. In writing the type and extent of the writing to be evaluated would need to be specified. Once the domain is specified, content validity depends on the ability to demonstrate that the items in fact represent that domain.

Two facets of content validity are found in the literature of MCT and in this literature are called curricular validity and more recently instructional validity (McClung, 1977). Neither terms will be found in the APA Standards, both fall under the more general rubric of content validity. However, because these terms are so widely used they need to be discussed here.

Curricular validity involves demonstrating that the test items measure content and behaviors that are part of the curriculum materials used in the schools. Curricular validity is seen as being important because if the competencies are required in order to receive a diploma, then those competencies must be a part of the school curriculum. If they are not part of the school curriculum then it would be unfair to deny a person a high school diploma because he did not posssess such competencies. In other words, if the competencies are not reflected in the curriculum then presumbly they are to be obtained through some sort of extra school experiences and therefore they should not be linked to the certification of the successful completion of secondary school.

Whatever model of implementation is adopted for a MCT program the certifying agency would almost undoubtedly have to show that the test had curricular validity. If the domain of interest is defined in terms of school specific skills then the test, unless set by the LEA, would have to measure competencies common to all of the curriculum material used throughout the state for the subject in question. Demonstrating curricular validity can become complicated when the test is given somewhere between grades 9 and 11, since a host of curricular materials have been employed by various districts over the grades prior to the grade at which competencies are being certified.

If the competencies are worded in terms of life skills, i.e., balancing a checkbook, filling out a job application - determination of both content and curricular validity is further complicated. These life skills essentially ask

the pupil to apply basic school skills to new situations. As we saw, the hope is that specific school skills like addition, reading, etc., in fact transfer to more general life skills like balancing a checkbook or reading a prescription label. Thus content/curricular validity may be quite hard to demonstrate for a test measuring life skills for two reasons. First, it may not be possible to define and delimit the domain of life skills. One could always ask whether a new skill falls in the domain. Second, such skills are not always common to the curricular materials used in the schools. Interest rates, checkbooks, tax forms, etc., were included in the 7-8 grade arithmetic books in the 40's and 50's but dropped out of sight in the 60's with the advent of modern math, the exception being courses like business arithmetic where such topics are still covered. Thus when an agency is interested in certifying "life skills" it may find it has a problem in specifying the domain of life skills and second, in showing that the behaviors in the specified domain are common to the curricula used by the various LEA's.

Instructional validity is a term coined by McClung (1977). He argues that it is not enough to demonstrate that the competencies are common to the various curricula; it must also be shown that an individual pupil received instruction directly related to the competencies measured. Such a demonstration could involve rather extensive record keeping. For example, teachers at all grades might be required to keep a log of the time spent teaching each competency for each pupil. The homework assignments, classroom tests, etc. for each pupil might also be used as indicators of "instructional validity."

"Instructional" validity is closely linked to the legal concept of "right to treatment" (Edmonds, 1977). In the area of mental health the courts have recently ruled that adequate treatment is the necessary condition for involuntary confinement. Using an analagous argument the court could rule that before an agency could deny a student a high school diploma because he/she lacked some set of minimal competencies the agency must show that the pupil - who was compelled under law to attend school until he/she reached age 16 - received "adequate" treatment. If a decision were to be made involving the principal of adequate treatment then McClung's "instructional" validity would become very important.

If the competencies are limited to scholastic skills then "instructional" validity would not be hard to demonstrate. However, if the competencies are framed in terms of life skills then instructional validity could pose a problem unless special courses, or blocks of time, were specifically devoted to teaching the application of basic skills to life situations. It should be noted that demonstration of instructional validity does not necessarily mean that most students will master the competencies only that they received instruction.

In fact, "life skills" might be taught but certain populations of students might not master them. NAEP found in their assessment of functional literacy that 17-year olds had the poorest performance on exercises which involved the use of reference materials; reading forms to identify the correct amount of coverage for bodily injury liability on an automobile insurance policy; understanding the billing on a book club membership form; and understanding

a traffic ticket. It could be that 17-year olds were never taught these skills. We noted elsewhere that application skills like these are difficult to teach. Further, the domain of such skills is extremely large and instructional coverage would have to be limited to classes of skills (i.e., reading forms). In the adult world the number of forms that one might have to fill out, or instructions one might have to read, are endless. Therefore instructional validity would have to be limited to the application of a skill like reading to well defined classes of adult situations (i.e., filling out government forms) simply because the schools could not possibly cover all specific situations.

When the competencies are stated in terms of life skills, where the student is required to apply basic skills of literacy and numeracy to real life situations the validity of an indirect paper and pencil multiple choice measure of these skills becomes an issue. Examples of life skill competencies might include: ability to fill out a job application form correctly; complete a tax form or application for a loan; balance a checkbook; read the local newspaper with comprehension; write a business letter correctly; spell and apply the rules of grammar in written work; check on the accuracy of bills, sales slips, etc.; understand basic contracts for car loans, house mortgages, etc.; follow directions correctly on prescription labels, in cookbooks, appliances, etc.; understand the terms of warrantees; be able to use the public library or other town, state or federal offices; plan and prepare nourishing meals.

Competencies like these are most validly measured by the most direct means possible - situational or performance tests - to determine if the student can or cannot perform these activities. Even when the competencies are limited to traditional scholastic areas such as writing or listening, the validity of using indirect multiple choice tests to assess these skills is questionable. To frame multiple choice items that ask a student to identify an example of poor, or incorrect writing, is no substitute for the more expensive and direct measure of asking the student to write a paragraph or a business letter. In the case of writing and listening the relationship between such direct and indirect measures is very low and therefore the indirect measure should not be used.

When the number of individuals to be certified is large then direct measurement is often costly and time consuming. Because of these problems indirect measures are often substituted. This is most clearly illustrated in indirect attempts to measure writing skills through the use of objective multiple choice tests (Foley, 1971; Madaus and Rippey, 1966). However, indirect paper and pencil tests measuring knowledge about performance skills rather than the skill itself are not enough. Any indirect, or surrogate measure will have to be validated against direct performance measures before they can be used. If the relationship between the direct and indirect measures is low then the direct measure must be employed.

In the case of writing and listening skills direct performance measures will have to be used. Traditionally there has been difficulty in objectively

scoring writing tests. However procedures can be set up that will assure acceptable levels of inter and intra rater objectivity in scoring. These procedures will not be inexpensive but nonetheless must be used or the results will be challenged.

If the competencies to be certified are linked to adult functioning then the certifying agency may be required to demonstrate that MCT's have criterion validity. That is the agency may have to show that the MCT's are actually related to certain extra school activities.

If the area of employment testing is any indication then demonstrating criterion validity could prove to be a costly and frustrating task. First, a criterion measure of adult functioning would have to be developed; questions about the validity of this criterion measure would have to be dealt with even before one were to study its relationship to performance on the MCT's. There are any number of ways adults can compensate for some deficit, or lack of skill, and function effectively. The definition of a term like "effectively" is not an easy matter and ultimately is a judgmental, value decision. Another problem we alluded to earlier is that definitions of effective adult functioning can change with developments in the society. The need for new skills that we cannot anticipate can arise.

Problems of defining and measuring some criterion of effective adult functioning aside, long term predictive studies show weak relationships between scholastic or school achievement measures and success in various professions. Suffice it to say at this point that if the state decides to word

their MCT program in terms of adult functions, then the criterion validity of the certifying test(s) will become an issue. The implications of undertaking such a program need to be understood before any challenge to the test(s).

11. What are the issues related to the difficulty level of items on a minimal competency test?

Very simply, item difficulty is the percent of examinees who answer an item correctly. Thus the index of item difficulty (percent answering the item correctly) can range from 0 to 100; the higher the index the easier the item. In a traditional norm referenced test, the test constructor tries to obtain an average difficulty level across all items of about .625. Further, items which almost all candidates in the item screening phase of test development get correct or incorrect are discarded in order to arrive at the desired average item difficulty level. Such an average difficulty level helps to maximize test score variance which in turn helps to assure acceptable reliability coefficients. The effect of selecting items with this difficulty is a test which maximally differentiates between individuals and results in a more or less normal (bell shaped) score distribution.

In building a MCT, it becomes imperative that items are tried out beforehand to get estimates of their difficulty before the test is finalized. The reason for this is simple; as is the case with traditional norm-referenced tests, the items are the individual building blocks that eventually produce the desired score distribution. If the cut-score for pass-fail on a MCT is set at 90% of the items answered correctly, the items included in the test

should on average have a difficulty level equivalent to the cut-score. Thus, item difficulty cannot be separated from political considerations in setting the cut-score.

As we have mentioned earlier, one reason offered to explain why 36% of the students in Florida failed the math portion of that MCT while only 8% failed the communication skills portion is that the average item difficulty level on the former portion was considerably lower than on the latter. Glass (1978) points out that someone writing items to cover a given topic can produce some that turn out too hard and others that turn out too easy but which on inspection appear to be equivalent.

Glass (1978) provides an example of how item difficulty can affect the inferences one makes about competencies. Two questions in a New Jersey assessment were designed to test the objective, "adding single digit numbers." In one item the digits are arranged vertically, and in the other horizontally. Eighty-six percent of the third graders tested got the vertical item correct, but only 46% got the horizontal item correct. This example clearly illustrates how item format can contribute to difficulty and perhaps lead to incorrect inferences about the skill being measured.

There are other reasons for pre-testing items. Item analysis can reveal ambiguities in the item stems or distractors. Further it can reveal items which are differentially difficult for different groups of pupils. If the test makers believe that the level of achievement should <u>not</u> differ between two groups (say boys and girls), they can reject items which are more difficult

for one or the other of the groups. This procedure would result in a test in which members of both groups perform equally well on average.

12. Can competency and item pools be developed?

One method that might be used to construct MCT's for use by LEA's is for the Commonwealth to develop a bank or pool of competencies along with a bank or pool of items designed to measure those competencies. Techniques for developing such pools have been worked out and there are extant, large pools, or banks of statements of objectives for many curriculum areas. Generally in such a pool each objective is coded by subject, by topic within subject and by specific elements within a topic. The objective is further coded as to the behavior or skill desired - often the Taxonomy of Educational Objectives forms the basis for such coding. An example of objective item pools include those developed by the state of Alaska, New York State, the Instructional Objectives Exchange, the Multinomah County Intermediate District in Oregon, and the Sequoia Union High School District of Redwood, California.

Thus, if the state decided to build a bank or pool of competencies and items there are models to follow and cost estimates could be obtained. To begin such an undertaking the Commonwealth might buy into several existing pools. Unlike curriculum referenced objective pools, the construction of pools of competencies needed for graduation would have to involve representative groups of teachers, lay people, parents, minority group members, etc. The same holds true for item development. Participation by various concerned

publics would have to be built into the process. As we argued previously, court challenges could turn on citizen participation - particularly that of minority groups - in the development of both competencies and model items. Alternatively, one could have educators work with measurement people to develop the competencies and items and then have the products evaluated by various publics, through an opinion survey type procedure.

13. What are the costs of a minimal competency testing program?

With expenditures for education already amounting to approximately 8% of the nation's gross national product, and per pupil expenditures increasing by roughly 50% (adjusted for inflation) over a ten year period, the costs might well be a determining factor in both the design of a new program and in how it is implemented. Interestingly, even given the rising costs for education noted above, one 1976 poll indicated that 52% of the American public thinks too little money is being spent on education. Only 10% responded "too much", and the remaining 38% indicated the amount being spent was "about right." Lest one assume too quickly the American public is willing to translate these opinions into action, however, consider that in the decade 1965-1975, both the absolute number and the percentage of school bond referendums approved by the public dropped noticeably. About 75% of such bond proposals won in 1964-65, compared to less than half in 1974-75. Monies approved dropped from $2.5 billion to $1.2 billion in the same period. One implication of this is that the public may not be willing to contribute to increased local expenditures for schools but may prefer that the increases come from the state and federal level. Indeed, an informal poll of teachers, administrators,

parents, and students taken in conjunction with public meetings on Minimal Competency Testing conducted by the Bureau of Research and Assessment indicates that about 80% of the 464 respondents felt that local school districts should pay none or only some of the costs of a minimal competency program. Thus, not only might the absolute costs of a minimal competency program be important, but the source of funding may be an equally important consideration.

It may be fortunate that economies of scale indicate a state operated testing program would cost less in some respects than a decentralized one, while, at the same time, it places the responsibility for funding on the state. However, the resolution of the cost issue is unlikely to be solved quite this easily, because there are perceived undesirable and intangible costs associated with a centralized program which in turn can lessen the grass roots support for a MCT program. Such grass roots support is, naturally, vital to the success of any MC program.

There is a broad range of costs associated with any minimal competency testing program; obvious ones such as test development; more obscure ones such as possible legal costs in defending against law suits. In addition to the monetary costs, there are also intangible costs, such as teacher dissatisfaction or local level antagonism at receiving another mandate from above. The paragraphs that follow are divided into three sections. First, a brief explication of possible costs is presented under four general categories;

development, administration, consequences, and intangibles. Most of the costs identified will have to be met in one form or another regardless of the eventual format of the MCT program. Second, some of the factors which influence the magnitude of the costs are discussed. It will be noted that these factors affect different costs in different ways, so that their influence can be manipulated to some extent by balancing one cost against another. Third, some rough estimates of the dollar amounts associated with a MCT program will be provided.

Possible Costs

A. Developmental stage

The word "development" conjures up the notion of test development, but this is only the most obvious of several costs which may be incurred in developing a complete minimal competency testing program. The use of the word "development" also tends to be linked with the idea of one-time costs; however, most of these costs are ongoing, albeit at a reduced level, if the program is to be maintained and updated. At least six specific costs are associated with development of a MCT program:

1. Competency identification and codification. This process can be long and complex, involving widespread information gathering at the outset, and concerted efforts by teams of "experts" at refining the statement of competencies to be included in the program. Regardless of the route taken, competencies must be identified, operationalized, and disseminated hopefully with a spirit of consensus. All of this costs time and money.

2. Curriculum development/modification
   a. From the identified competencies to be tested, behavioral objectives need to be developed both as a basis for instruction and as a basis for testing.
   b. Curriculum in the form of guides, lesson plans, course outlines, or "planned course statements" similar to those developed in Oregon may have to be produced.

3. Public relations. At the very least the public must be informed of the nature and purposes of the program. The Oregon example illustrates that this may be accomplished along with curriculum development. Not only do the Oregon "planned course statements" serve as curriculum guides but as information for the public as well.

4. Inservice training. There may be a need to train staff in the administration and/or interpretation of the test results. In any event, staff time will be required for instruction and providing information regarding the implementation of the program.

5. Test development. Regardless of the manner of test development, local, state, or commercial, there are costs associated with it. If an already available test were to be used, these costs might be labeled procurement costs and would be very much an ongoing concern. Subsumed in test development costs are the costs of pilot-testing which would be required regardless of the test to be used. If a new test must be developed each year there will be yearly test development costs.

6. Test validation. As an adjunct to development costs, there may be substantial costs associated with test validation, especially in the case of indirect, paper and pencil measures. Even if development costs are circumvented by using already available standardized tests, one could still incur costs in validating the test for a particular use and situation.

B. Administration

As with costs in the development stage, there are additional costs associated with the administration of a MCT program than those associated with the mere administration of tests. Clearly, most of the costs involved at this stage are of an ongoing nature and continue at a relatively constant level year after year. Six examples of such costs follow:

1. Test administration including scoring. These costs obviously vary widely depending on many factors, type of test, frequency of administration, number of grade levels involved, machine or hand scoring, etc. It should be noted that there is a possibility for cost benefits here. For example, if tests could be administered to large groups of students by clerical personnel, the teaching staff may be free for other tasks during testing periods.

2. Record keeping. Virtually all the local districts, and states which have experiences with competency testing programs, report record keeping as a major problem and concern. It is conceivable that new staff may be required here depending on the complexity of records desired. As we noted previously, demonstrating instructional validity may require quite extensive long term record keeping.

3. Enforcement. A strong word, perhaps, but compliance with a testing program is by no means automatic. There is obviously a legal concern here; the more important the decisions to be made on the basis of competency tests, the more important compliance becomes. There may have to be a vehicle developed for enforcing compliance.

4. Monitoring. Apart from enforcement concerns, the regulating agency, state or local, has an obvious need to monitor the operation of a testing program and insure its responsiveness to changing concerns.

5. Provisions for special needs and bilingual students. If the decision is made, as seems likely, that special tests are needed for some students in these categories, then alternative methods of assessing their competencies will have to be employed. Almost certainly, these methods will be more expensive on a per pupil basis than the regular tests especially if the regular MCT is the paper and pencil variety.

6. Additional staff. Depending on the nature of the program, there may be a need for new staff for test development, interpretation, and in service needs.

C. Consequences

The costs of a MCT program do not stop once the test scores are in. In fact, one of the major cost concerns expressed by those working with MCT programs is the cost of compensatory/remedial programs for those

pupils who fail the test(s). Several of the costs associated with the period after MC testing are briefly discussed below.

1. Remediation/compensatory programs: Here again there is an underlying legal concern. To the extent that important educational decisions will be made on the basis of minimal competency tests, there is a legal responsibility to provide not only the instruction in the first instance, but remediation for those pupils who fail. Legal issues aside, presumably the goal of minimal competency programs is not to weed children out and save the fittest, but rather to insure that as many children as possible attain the competencies. This goal will almost certainly require remedial programs. It is entirely possible that additional staff may be required to provide such remedial instruction, particularly at the secondary level.

2. Legal: Perhaps unfortunately, it is likely that lawsuits will grow out of minimal competency programs, possibly in the form of so-called educational malpractice suits. Such suits involve obvious direct defense costs as well as indirect costs . In the latter category, for example, are record keeping costs referred to above. Accurate individual records of when specific competencies were taught may be legal money in the bank.

3. Reallocation of resources:

   a. From the gifted and average student to the low achiever. This tendency already associated with the implementation of Chapter 766 may be aggravated by minimal competency programs, especially to the extent programs are locally funded.

   b. From enrichment courses to core or basic courses. It seems almost unavoidable that if basics (as defined by the accepted competencies) are to be the focus of new programs, other courses, e.g., art, music, foreign languages, social studies, science, may experience a decrease in financial support.

4. Dissemination of information: The public will want to know the results of minimal competency testing. Information will have to go out on two levels: to the public at large about the overall performance and to individual parents about their children's performance. The information required will likely require more than merely reporting scores and will involve some guidance on how the results are to be construed and interpreted.

D. Intangibles

Costs in this category are by their very nature difficult to identify, to anticipate, to describe, and to quantify. Yet they are very real and, perhaps more than the monetary costs, can scuttle a program. A minimal competency program is doomed to failure without the support of the clients and staff of the public schools. Fortunately, several of the costs discussed below might be better characterized as risks. To the extent these risks are identified and anticipated, their cost can be held to a minimum.

1. State-local district polarization. As noted above, economics of scale may indicate a centralized testing program as the economical way to go. Furthermore, the general public probably favors funding flowing from somewhere other than the local level. However, if a minimal competency program is conceived, developed, and implemented from the state level, there may be a real risk of lack of local support for the program. Organizational theory posits that change is most likely to succeed when those individuals who are most affected by the change participate in defining the problem, setting the goals and standards, and developing alternative strategies for implementation.

2. Local administration-teaching staff polarization. The same concern discussed in (1) above may be a problem at the local level as well depending on whether the teaching staff sense "ownership" of the program or whether they sense it as forced upon them from school administration. The risks here are of teacher dissatisfaction, antipathy, and complaints of being overburdened. Unless teachers are fully supportive of the program it might well become a collective bargaining issue as well.

3. Public dissatisfaction with school results. Although there is apparently a public call for competency testing, such a program will not be a panacea for whatever ills have generated the call. Public expectations, however, might be very high, and if the results do not live up to expectations immediately, the public will likely be unhappy.

4. Antagonism from parents whose children fail. As we noted earlier there is a very real philosophical and legal question over who should be held accountable for the attainment of competencies;

the child, the teacher, the school, the school system, etc. One thing is sure; most parents will hold someone other than their children accountable. Therefore, parents of children who fail form a public with a real potential for antagonism toward the schools.

5. Self-concept of pupils who fail. The whole notion of social promotion now called into question by some competency testing proponents is predicated on avoiding stigmatizing children. A minimal competency test with a clearly defined cutoff point holds the possibility of labeling children to their detriment. As with certain other costs, the more important the educational decisions to be made based on a competency test, the greater the risk involved in damaging a child's self-concept.

6. Lowering standards to the minimal competencies. Currently operating programs tied to graduation generally require performance at an eighth grade level in order to pass. There is a danger that the level of competency required by the test becomes maximal, rather than minimal; the ceiling rather than the floor for educators and pupils.

7. Less effort put into non-core subjects. In addition to the monetary reallocations noted above, non-core courses may also suffer from intangible losses in pupil, staff, and public attention and support.

8. How will private/parochial school pupils be handled. Will they be required to pass MCT's in order to receive a high school diploma? If the answer is no, there may be repercussions from parents in both the public and private sectors.

Factors Influencing the Magnitude of Costs

It should be obvious that the costs of a MCT program vary in their impact. Many of the costs identified above may be either negligible or prohibitively expensive. For example, the cost of administering a test has a recognized range of from 15 cents per pupil to $13 per pupil, the former being a figure

cited by Denver authorities as the cost of administering their competency tests, the latter a figure associated with a secure test administered by a contractor for high school certification. Policy makers should be aware that the general magnitude of costs is largely determined a priori by decisions they make on the purposes and design of their program. In this section we shall briefly discuss some of the decisions which impact greatly on costs. The decisions concern the following factors:

1. The model chosen
2. Type of test
3. Number of competencies to be tested
4. Nature of the decisions to be made based on the testing program
5. The per cent of pupils one is willing to fail
6. The per cent of pupils who are expected to master the competencies
7. The difficulty of the competencies
8. Frequency of testing.

The model one chooses to implement for a MCT program influences costs in several different ways. The key feature of the model which impinges on costs is the degree of state involvement. In general, it would appear that some tangible costs decrease with state involvement due to economics of scale and centralization of tasks, while some intangible costs may rise due to weakened local support for a state-developed and administered program. This latter possibility, if foreseen by the state, may require increased expenditures for public relations aimed at obtaining local support. It is

clear that centralization could decrease such tangible costs as yearly test development, competency identification, curriculum development, public relations aimed at the public at large, and test administration and scoring. However, such intangibles as state-local district polarization and teacher dissatisfaction may be aggravated by a high degree of state control.

The type of test one chooses has obvious implications for test development, administration and validation costs. The choice of an available standardized test avoids development costs, may minimize administration costs, but increase validation costs. On the other hand, the use of a direct, performance-oriented, situation-specific measure may make the validation issue moot (at least for content and construct validity) but be expensive in terms of administrative costs. A performance test may also be much easier to "sell" to test-skeptical publics than a multiple choice test.

The number of competencies to be tested, and the number of items per competency, have a direct bearing on test development costs, especially for paper and pencil tests. The development costs of such tests are usually figured on a per item basis. Many of the school districts in Oregon, where local control is the rule, began with lists of from 150 to 200 competencies, but reduced this to an average of 33 for reasons of cost and ease of development.

The nature of the educational decisions that are based on competency testing has an across the board influence on costs affecting virtually every area of the program. In general, the more important the decisions to be made, the higher the cost. Several examples may serve to illustrate the

point. First, when important decisions are to be made like awarding a diploma or not, test development costs would likely be much greater because the test will have to be technically above reproach. Longer tests would be required to increase reliability, and there would be increased likelihood, even certainty, that a contractor would be required for test development until the state could build its own test development bureau. Second, curriculum development, remediation, and legal costs would likely increase with the degree of importance of the decisions. Third, validation costs are directly related to the importance of decisions being made. School personnel, the public and the courts will want to be convinced a test was valid for the specific situation and purpose for which it was used.

The failure rate one is willing to accept directly influences remediation costs. Quite simply, remediation costs increase with the failure rate. Additionally, high failure rates, especially beyond some unknown threshhold point, would likely increase public dissatisfaction with the schools and would involve increased public relations costs in an attempt to offset such dissatisfaction.

The per cent of the pupils one would expect to achieve mastery of competencies would also affect costs. Here the policy concern is not with how to juggle a cutoff point, but once a cutoff point has been determined, to decide how many pupils one will work with to bring to that mastery level. Obviously the higher this percentage, the higher the costs.

The difficulty of the competencies one selects also influences costs. There is no doubt that some competencies are easier to acquire than others. To the extent difficult competencies are selected instructional costs will increase.

Finally, the frequency of testing influences costs in several ways. The most obvious of these is that there is a direct relationship between frequency of test administration and costs. Additionally, both initial and ongoing development costs may increase significantly with increased frequency of testing. For frequent testing, large item pools and parallel forms would likely have to be developed.

The discussion above is by no means intended to be an exhaustive one on the variables affecting costs. It is meant rather to indicate the complexity of the cost issue and the fact that the costs are highly variable.

Cost Estimates

Many, if not most, of the costs identified in this paper are difficult, perhaps impossible to estimate except within the context of a very specific program. Some of the intangible costs may not even be quantifiable. Program developers will have to anticipate the way in which different variables might affect cost in general and then weigh the cost considerations in light of their program goals and design. Disclaimers aside, a rough idea of ranges of costs can be made in general terms for several areas associated with a MCT program.

1. Test development. Costs run from $25 to $210 per item with an average cost of $100. Assume a test covering 30 competencies with 5 items per competency. In order to obtain 150 items after the pilot testing, roughly 450 items are needed. Thus, one is talking in the neighborhood of $45,000 for test development. Since the tests might not be able to be used again this could be a yearly cost.

2. Test administration and scoring. This is an area subject to wide variation. Cited above are figures ranging from 15 cents to $13.00 per pupil per administration. One factor which clearly affects this cost is the type of test one is using. For example, administering a commercially available standardized test would likely run about $2.00 per pupil. Administering, monitoring, scoring and interpreting performance tests could cost even more. Our personal experience with administering and scoring writing samples indicates that the cost for each paper rated on only one criterion (i.e., a general overall rating of the paper's quality) would be around $6.00. If a greater number of ratings were desired this cost would increase but not dramatically. For example, 15 criteria (i.e., separate components of writing) could be rated at a cost of about $7.50 per student.

3. Remediation. Select a percentage one expects to fail, determine the time to be spent on remediation, class size in remedial programs, and teacher salary to create a cost estimate. For example, if each child needing remediation attends 3 classes per week; a teacher teaches 21 classes per week; 210 students need remediation; and the average class size is 10 students; then 3 teachers can cover all the classes. The cost would thus be 3 teacher's salaries plus all the ancillary costs associated with these salaries (e.g., fringe benefits) and a remedial program (e.g., materials, overhead).

4. Bureaucratic costs. Anderson (1977) reports state bureaucratic costs ranging from $65,000 to $945,000 per year. Local control may multiply these costs.

5. Curriculum development. Determine the number of teachers required to revise curriculum, the time required, and average yearly salary. For example, in one district, three English teachers work for three weeks during the summer, average annual salary $12,000; cost approximately $3000.

The costs are many, possible variation almost limitless, estimates difficult to come by; but the cost issue itself may be of overriding importance.

Summary

The following is a listing of the major points presented in the section (Part II) of the report.

1. The MCT movement has grown partially in reaction to a set of beliefs that schools are neglecting the 3 R's and that scholastic "standards" are declining.

2. The language of MCT programs focuses on the ends of the programs and draws people away from a consideration of problems associated with the means necessary to attain such ends.

3. Minimal competencies should be justified in terms of community consensus on their importance and not in terms of their being essential life skills or necessary for adult functioning.

4. Most programs are structured in such a way that the pupil is the only individual that is directly held accountable for failure to demonstrate minimal competencies. By testing from grade 8 onward, poor performance can only be attributed to the pupil or the "system." It would be quite unfair to try to hold an individual teacher accountable.

5. MCT programs are apt to impact most heavily on that segment of the student population which is weakest scholastically.

6. We should not wait until grade 8 to identify pupils needing help in the minimal competencies.

7. Minimal competencies can be classified as "enabler skills" or "life skills." The former are basic scholastic skills associated with the three R's; the latter involve the application of enabler skills to life situations like balancing a checkbook, computing the unit price of an item, filling out income tax forms, etc.

8. In order to include a competency in a minimal competency testing program, there should be a strong belief that it can be taught to, and learned by, all pupils given sufficient time, resources, etc. Enabler skills fit the requirement better than do life skills.

9. Affective competencies should not be included as part of a MCT program tied to graduation.

10. Competencies selected should be amenable to remediation when necessary, and a remedial program must be developed as an integral part of a minimal competency program.

11. A curriculum must either exist or be developed, designed to foster the selected competencies.

12. The number of competencies selected should be kept small and should be drawn from basic cognitive areas that are presently part of the school curriculum.

13. Minimal competencies should not be allowed to become maximal in the sense that they are emphasized to the neglect of other areas of the curriculum.

14. Various publics need to be involved in stating or selecting the competencies to be included in the MCT program. The participation of parents, teachers, employers and representatives of the poor and minority groups should be actively involved in the selection process.

15. Testing should occur early and often so that multiple opportunities to sit for the test are afforded to pupils.

16. Pupils with a high probability of failing the minimal competency test should be identified as early as possible so that remediation can begin promptly.

17. A danger associated with identifying pupils with a high probability of failing the minimal competency test(s) is that they will be pejoratively labeled, setting in motion a self-fulfilling prophecy of failure.

18. Communities with high minority populations must be careful that by identifying and remediating pupils at risk they do not re-segregate pupils.

19. A pupil needs to pass the MCT only once in order to be certified. Once the test(s) have been passed pupils should not be required to demonstrate the competencies again at a later date in order to graduate.

20. Most handicapped pupils should be treated as individual cases in determining how, or if, they should be required to demonstrate all or a subset of the required competencies.

21. Non-English speaking students have to have sufficient opportunities to learn competencies requiring standard English.

22. Minimal competency tests are set in standard English; if poor minority children do poorly on such tests it may be attributed to the instructional/learning process rather than to test unfairness.

23. Programs of remediation need to be a required component of any minimal competency program.

24. Secondary teachers generally are not trained to provide remedial instruction in basic skills so that a number may have to be re-trained or elementary staff shifted to work with secondary pupils.

25. Policy makers need to explore the impact of a minimal competency program on Title I and Chapter 766.

26. A MCT program may not result in an increase in remedial services over present levels but may re-focus those services in terms of record keeping and testing in order that a school can document that any student in need has received the required services.

27. In designing a remediation program for MCT programs, we need to be able to: identify pupils in need; focus on the specific factors underlying the difficulty; and design an instructional program that will help overcome the difficulty.

28. The effects of a MCT program on the school curriculum are a function of the perceived importance of the tests, the extent to which the competencies are already a part of the curriculum, and the amount of lead time before the required program takes effect.

29. A choice must be made between a norm referenced and criterion referenced standard for passing a minimal competency test(s).

30. If a norm referenced standard is used the norms must be based on a high school population and should be set at a grade level below twelve.

31. A standard based on answering a fixed percentage of items correctly is criterion referenced in the sense that interpretation does not depend on the performance of other pupils. However, such a test may not be criterion referenced in the sense that performance on the test is linked to a specific skill domain.

32. Setting a cut-score, or standard, be it norm or criterion referenced, is ultimately a value-laden, judgmental process. There are no agreed upon statistical procedures that remove human judgment from setting standards.

33. A decision has to be made on whether or not each selected competency will be certified or whether some overall total score will be used for certification.

34. It is not possible or desirable to try to use the same test year after year. Minimal competency test(s) should be secure only up to the time they are administered and then should be made available for public scrutiny.

35. Any minimal competency test must be content valid. It must be able to be shown that the test samples content and behaviors which are part of the curriculum used in schools. Demonstrating content or curriculum validity may be difficult if the competencies are defined in terms of life skills since such skills are not always explicitly covered in the curriculum.

36. "Instructional validity" would require an agency to show that the skills measured by their minimal competency test(s) were actually taught to pupils. This could require extensive record keeping for each pupil.

37. Indirect measures of a particular skill must be validated against a direct measure of that skill.

38. If the minimal competencies are linked to "effective" adult functioning, then the minimal competency test(s) will have to possess criterion validity.

39. Items to be included in a minimal competency test should be tried out beforehand to determine their approximate level of difficulty.

40. The Commonwealth might build a pool or bank of competencies and items that LEA's could use to design their own minimal competency test(s). Various concerned publics should be involved in the development of such pools.

41. There are a wide range of direct and indirect costs associated with the development and maintenance of a MCT program which must be clearly anticipated in advance so that adequate funding can be provided.

PART III: COMPARISON OF SELECTED MODELS

The previous section overviewed a number of issues and concerns that may arise in the conceptualization or implementation of any minimal competency program. Thus, for example, concerns about purpose, need, and language are important regardless of the specific features of any proposed program. Similarly the potential problems associated with teaching, testing, remediation, and setting standards face the implementors of all minimal competency programs. Finally, given the present climate, cost and legal factors must be attended to by program planners. The issues discussed in the previous section, therefore, cut across varied programs.

However, some of these issues assume increased importance according to the specific plan or procedures any given minimal competency program chooses to adopt. Thus, for example, the issue of state control of the curriculum - always an important concern in thinking about minimal competency programs - may warrant greater concern when a program is conceived and carried out by state agencies than when it is subject to local district control. On the other hand, costs for the program likely will be greater if each district must carry out its own program rather than capitalizing upon potential economies of scale available through a centralized state conducted program. Moreover, when the minimal competency program is tied directly to graduation certification, many of the issues noted take on a heightened importance.

In this section of the report we will describe three models for the implementation of minimal competency programs which were selected by the Policy Subcommittee for particular consideration. Within each model we shall describe the problems and consequences most likely to arise if that model were adopted. Finally, we shall summarize the discussion by comparing the three models in terms of common and idiosyncratic trouble spots.

As noted above, there are some issues that will not be included in this section. For example, it was argued that an essential characteristic of any minimal competency approach is that questions regarding the purpose of the program and who the program is intended to hold accountable for learning be resolved. We shall assume that however these issues are settled, they apply similarly across the three models to be compared here.

The three models to be discussed were selected from a much larger array of possible models under which a minimal competency program might be implemented. To identify this larger set of models, the authors specified three central functions or activities in minimal competency programs:

1. Specifying or selecting the competencies
2. Setting or selecting the testing procedures used to examine mastery of the competencies
3. Defining the criteria or cut-off points which define acceptable performance on the tests.

We then asked, "What agencies or mechanisms might assume control of carrying out these functions?" Four possible mechanisms were selected:

1. the state, through some existing or to-be-created agency
2. the local school district
3. the local district selecting from state approved lists or options
4. the local district subject to state approval

The three functions and four control mechanisms were then crossed in all possible combinations to produce 64 different implementation models. Some of these models did not seem viable and eight of the 64 models were selected for presentation to the Policy Subcommittee. In general, these models represented models similar to those used in other states. The Policy Committee selected three models for more detailed comparison and description.

The three selected models were: 1) state control of specifying competencies, setting tests, and defining passing standards; 2) local district selecting from state approved options for competencies, tests, and standards; and 3) state specified competencies but local district selection of testing procedures and standards subject to state approval. In addition, for the third model, the Policy Subcommittee asked that a comparison be provided between a state or a state-endorsed agency setting minimal competencies and the state allowing local districts to set competencies subject to state approval.

The following sections will consider each of these models in turn, providing a description of the model, implementation concerns, and needed

resources. The critical policy issues which might arise if each model were implemented will also be discussed. In particular, the models will be discussed and compared in terms of such issues as: concern over local control of education, impact on the curriculum, remediation, establishing standards, measurement issues, legal issues, and cost. Obviously, it is impossible to elaborate all potential scenarios which might arise in each of the models presented. Instead, we shall endeavor to highlight many of the administrative concerns for each model and describe the most likely consequences resulting from each model's adoption.

Model 1: State responsibility for specifying minimal competencies, establishing testing procedures and setting standards

Under this model, a state or state-sanctioned agency assumes administrative responsibility for the minimal competency program, including the specification of competencies, the determination of testing procedures and the setting of the standard of performance which will indicate minimal competence. In essence, this model is the most centrally controlled of the models to be discussed, insofar as options within the program are limited. That is, the program encompassed under Model 1 is a state-wide program in which all districts and pupils within districts are responsible for the same competencies, tested by the same procedures and evaluated in terms of the same standard.

A central concern in this model is, of course, "what agency or agencies should assume responsibility for carrying out these activities?" There are

three logical possibilities. First, an existing state agency, the Bureau of Research and Assessment for example, might be charged with the responsibility for administering the program. Second, a newly created agency - a Massachusetts Minimal Competence Authority analogous to the Massachusetts Turnpike Authority - might be given responsibility for the program. Third, the entire program might be contracted out to a private agency or testing firm. It is also conceivable, of course, that these different groups could each administer one or more of the different program components, although this arrangement could complicate administrative responsiveness.

In deciding upon the agency which should administer the minimal competency program under Model 1, a number of concerns should be considered. First, as noted in the prior section of this report, the issues of minimal competency are complex, both from a value and a technical standpoint. It is essential then, that whatever administrative arrangement is selected, it have or have access to the necessary expertise to carry out the program. Although this recommendation sounds both obvious and simplistic, the current ferment in Florida over the results of this year's minimal competency testing points up the need for technical expertise in conducting a program. Second, the administrative agency must be acceptable to various special interest groups around the state. In this sense, a newly created agency might have the advantage of political viability state wide, whereas an existing state bureaucracy might be tainted simply because it is a state bureaucracy. It is not clear how an external, contracted agency would be perceived. The longer a bureaucracy or agency

has been in existence, the less neutral are the publics' attitudes towards it likely to be. Unless the administrative agency starts off with a reasonable amount of political viability across the state, implementation and conformity to the rulings of the agency may be difficult to obtain.

Finally, cost and time factors are important in deciding upon what agency will administer the state-wide minimal competency program. These factors should be balanced against the two factors just discussed. Creation of a new agency to direct the program may be appealing in terms of the likelihood of eliciting broad based cooperation, but setting up such an agency probably will be time consuming and expensive, especially if that agency is charged with the responsibility of carrying out the more technical aspects of minimal competency testing. However, selecting an existing agency to administer the program may increase the need for regulatory, legal and inforcement costs, particularly if districts balk over the viability of that agency controlling the program. Note also that diverting an existing agency from one task to another, while appearing cost effective, might not be in the long run. Some states have wiped out their Statewide Assessment Programs and diverted personnel to newly established minimal competency programs. While this strategy may be cost efficient in the short term, it means that five years from now those states will have lost the ability to examine the effect of minimal competency programs on students and the wider curriculum - unless, of course, some form of statewide assessment is reinstituted. At any rate, it is fairly evident that under a state-administered model, a substantial bureaucracy is required to handle the tasks of testing, test construction, record keeping, test

administration, certification, monitoring, responding to complaints, etc. In selecting an agency to administer the minimal competency program, the three concerns of expertise, public acceptance and cost must be carefully considered, both individually and as they interact.

In Model 1, a single set of state approved competencies exists. In order to obtain credibility and support for the program, it seems important that the definition of the desired competencies not be left solely to the discretion of the administrative agency, but rather be based upon wide ranging input from various interests and groups in the community. Thus, the definition of the competencies - and the direct responsibility of the administrative agency in defining competencies - should be separated from test construction and standard setting. The latter two tasks may be carried out largely by the administrative agency, while the former should be done through more broad based community participation; definition should be separate from control.

The single, state approved set of tests of minimal competency in Model 1 pose some additional administrative problems. As noted previously, it will probably be difficult to maintain the security of a given set of tests once they are administered. This likelihood will mean that new tests will have to be constructed frequently. Moreover, in a state-wide program, especially one which relates performance to high school graduation, it will be important to standardize many facets of the test administration process. For example, testing probably should take place on the same day or days state-wide in order to avoid problems of the tests becoming available to some groups prior to

their being tested. This requirement in turn implies the need for some kind of bureaucracy to insure uniform test administration conditions across school districts. Mechanisms for distributing, scoring, and returning the tests and test results will also be required.

The above issues reflect general concerns in implementing Model 1. We may now turn to a discussion of the likely potential impact of Model 1 on a number of areas. In essence, the question we seek to answer in each area is "If Model 1 is adopted and implemented, what are its most likely effects in this area?"

Issues of State versus Local Control

In any model in which there is a single, state approved and mandated set of competencies, testing procedures and standards, a tension between state versus local control of education is likely to arise. There are a number of benefits which may be associated with a minimal competency testing program such as that envisioned under Model 1. First, competence would tend to have a fairly uniform meaning across the state. With all pupils held to the same competencies and standard, there would be less variation in what a high school diploma means than is currently the case, although a great amount of variation would still exist. For those who perceive a need to bring uniformity and meaning to the diploma, Model 1 would likely provide a major step in this direction. Note that Model 1 sets equal expectations for students across the state, but that equal expectations do not imply equal opportunity, equal achievement, etc. Second, a state administered program similar to

Model 1 may have greater credence in the eyes of the public than a series of district level programs with different competencies, testing procedures, and passing standards. The possibility of some districts being labeled "easy" locales to obtain a diploma, thereby lessening the political validity of the minimal competency program in general, is reduced if all pupils in all districts must demonstrate similar competencies on similar tests to similar levels of proficiency. In the public's eye, there may be something "fairer" and "more rigorous" in having all pupils held to the same standards than if the standards vary from district to district. Centralized, state control of the minimal competency program therefore presents certain advantages.

However, these advantages must be considered in light of some disadvantages. A state administered model such as Model 1 is likely to take least account of district to district diversities which exist currently in the state. Not all school districts are alike and differences between districts are likely to be ignored under Model 1. Districts have different needs, types of pupils, teachers, resources, problems and priorities. What will be the impact on this diversity of a program which requires pupils across districts to adhere to identical competencies, test procedures, and standards? What will be the impact on those districts which have already begun to institute minimal competency programs in response to their own local needs?

Given that the state controlled program is tied to graduation for pupils, it seems reasonable to expect that the competencies stressed will assume major importance for school districts. While this may not necessarily be

bad - pupils will at least be receiving instruction in the desired competencies - it is likely that arguments about a state curriculum and infringement on local control will ensue. And in one sense these arguments will be valid, since the state, through its control of the minimal competency program tied to graduation certification, potentially will exert a strong influence on the curriculum and instruction at the local district level. Note that this influence is not general, as in many present state programs, but will be directed to particular competencies within particular subject areas. Any state administered minimal competency program such as Model 1 inevitably will raise concern over infringement on local control and the propagation of a state-wide curriculum.

Impact on Curriculum

The impact of Model 1 on local school curricula will depend in part upon the extent to which the state-wide competencies currently exist in local curricula. Potentially, however, Model 1 could exert a strong influence on local school district curricula for three reasons. First, local districts will not be in control of the testing and certification of their pupils and thus may be prompted to make an anticipatory response to an externally imposed, unknown examination. Second, the impact of the examinations for students, coupled with the lack of district control, will likely lead to an increased emphasis on the stated minimal competencies in instruction. Finally, the attendant publicity which undoubtedly will be associated with the implementation of a state-wide minimal competency program will increase parental pressure on the school district for satisfactory pupil performance on the examinations. A district's most logical

response to this uncertainty, lack of control and parental pressure will be to insure that the state-wide competencies assume a central role in classroom instruction. A move to a state-wide basic skills curriculum, an increased stress on basics, a reduction of emphasis in other areas, and conformity across districts are very real possibilities in a context in which Model 1 is in operation. Moreover, if enough pupils are failed or denied a high school diploma in the early years of a minimal competency program, there will be increased pressures on local curricula to conform to the state approved competencies.

These pressures are likely to be particularly strong for such skills as are involved in the communication competencies currently developed. Verbal and auditory communications are not now stressed explicitly in most school curricula. If certification for a high school diploma is dependent on demonstrated minimal competence in these areas, it is to be expected that programs and courses emphasizing these skills will be inserted into local school curricula. Given a fixed daily time period for education, some other areas of the curriculum are likely to be deemphasized to permit instruction in communications skills.

Potentially then, for the reasons cited above, Model 1 may have a large impact on local school curricula. The issue of the impact on curricula should be considered within the context of two factors. The first, alluded to in the prior section, is the desirability of local options and control in light of local needs as opposed to the desirability of a state-wide basics skills curriculum

which might result from implementation of Model 1. The second issue concerns the phase-in time necessary for localities to make the necessary changes in their curricula implied in the adoption of Model 1. The amount of time, energy and resources available to revamp existing curricula in light of state-mandated competencies will vary across districts and should be considered prior to tying minimal competency test performance to high school graduation.

Remediation

Regardless of what minimal competency program model is adopted, it is clear that remedial activities should be provided for pupils who fail to demonstrate minimal competence. Within the context of Model 1, however, there are two basic issues associated with remedial services. The first issue is who will pay for such services, an issue to be considered in a succeeding section. The second issue is the level at which remediation should be administered and provided.

It seems unlikely that the state agency in charge of the minimal competency program would wish to take on the direct task of providing remedial services in each local district. Doing so would require a gigantic bureaucracy. What seems more likely is that the state will leave the problem of remediation up to teachers and specialists in the local district. While this is a wise and probably efficient strategy, two potential consequences of this approach must be considered.

First, given state-wide competencies and the likelihood that many districts will have to restructure their curricula to meet these competencies, it is to

be expected that resources within districts will be reoriented and reallocated. Efforts to remediate skills, knowledge, etc. which are not currently emphasized at the district level will require a redistribution of funds and personnel from current activities. The pressure for remediation of the state approved competencies can therefore impact on local district discretionary funds and, consequently, influence the nature and types of programs a district can continue to provide. Such a result is a consequence of a local district having to provide remedial services in areas it does not emphasize in its curricula but which the state requires for high school graduation.

Second, even though the state administering agency will not be responsible directly for providing remedial services to pupils at the local level, it seems likely that the state agency will want to monitor the local districts to be sure that its desires and policies regarding remediation are being acceded to at the local level. The state will want to be certain that funds earmarked for remediation are spent on remediation, that remedial funds are not being concentrated on certain groups (for example, those with the greatest likelihood of passing the competency test if given remediation) to the exclusion of other groups ("hard core" learning problems), etc. In Model 1, one consequence of remedial activities, even though these are carried out at the local level will be the need for personnel and mechanisms for monitoring local districts' remedial procedures.

Standards

The existence of a single set of state-approved standards in the various basic skill areas leads to a number of consequences. First, the problem of

inter-district reciprocity which can occur when districts use different standards is overcome. Second, the high school diploma will have a more consistent meaning across districts. Third, the credibility of the high school diploma and the minimal competency testing program itself will likely be enhanced by the use of a single state-wide set of competency standards. These consequences are fairly obvious and have been described previously.

Two additional, perhaps less obvious, consequences can ensue from the use of a single set of state wide standards of performance. It seems more likely that there will be challenges to state-wide standards on the basis of their unfairness to different districts or the standards' inherent bias against certain groups. For example is it "fair" - and if so with how long a phase in period - to compare urban pupils to suburban pupils on the same standards? Do resource, pupil, and other differences between districts require either different standards of performance or adequate phase in or gear up periods? Will disproportionate failure rates for different racial or ethnic groups lead to challenges to the state approved standards? The possibility of such challenges is a very real one in the context of Model 1. This, of course, is not to indicate that such challenges will be sustained, simply that they will occur and may move discussion of minimal competency programs from the educational to the emotional or political arena.

Additionally, the availability of district wide pass rates on the same set of standards will lead inevitably to inter district comparisons of performance. Such comparisons will undoubtedly put many pressures on local school districts.

Some of these pressures may be beneficial in light of the purposes of the minimal competency program. Others however, may not be beneficial, particularly when teachers, administrators, or school committee members are singled out or indicted for the performance of their district relative to other districts.

Measurement Issues

The consequences of Model 1 in the area of measurement are principally in the following four domains: administration and scoring, test validity, diagnosis versus certification, and comparability across districts. Model 1 calls for the administration and scoring of state-wide minimal competency tests. Due to the importance of the tests, assuming performance on them is tied to high school graduation, it will be essential for the state agency in charge of the program to exercise considerable care in standardizing the test administration and scoring procedures across local districts. Test security must be maintained until testing time. Pupils across the state's local districts must be tested at the same time and under the same conditions to prevent knowledge of the test and test content from spilling over from tested to untested districts. A "testing day or days" similar to that used for administration of the College Board tests, with similar administrative control and standardization across sites, will be necessary under Model 1, if only to insure "fair" comparisons across pupils and to reduce the likelihood of legal action associated with inequities in the minimal competency test administration. The necessity for such standardization of testing procedures across sites will require, at

the least, some form of monitoring or policing agency. Note the difference involved in administering minimal competency tests with important consequences tied to pupils' performance and administering a state-wide assessment device to a sample of pupils in a school. The former situation requires much more stringent control over the testing procedures than the latter and mounting a state-wide minimal competency program will be more involved than carrying out a survey of pupil achievement with no individual pupil sanctions applied.

Earlier in this report we overviewed a variety of test validity concerns as they relate to minimal competency tests. It is most likely that challenges to the minimal competency program will come either in the context of the application of a uniform state-wide standard for certifying competence or in the context of the validity of the tests used to assess competence. Obvious validity problems ensue from the use of indirect measurements ("pick the misspelled word") instead of more valid direct measurements ("spell the word 'competency'") and from the use of timed testing situations for tasks which need not necessarily be timed (e.g., fill out an income tax form in 5 minutes versus fill out an income tax form within two days). However, other validity issues will arise as a result of the use of a single, state-wide set of competency tests to assess what inevitably will be curricular diversity across districts. While many of the consequences associated with the lack of match between competency tests and district curricula emphasis may be overcome with adequate phase-in periods, it seems likely that in the context of state-wide tests and state-wide standards which impact directly on the individual pupil,

the state or state agency in charge of the minimal competency program will have to engage in some monitoring or policing of local districts to insure that the content of the state wide tests is reflected in local curricula and instruction. Of course, as noted above, one consequence of such policing and monitoring is the threat of moving towards a state-wide curriculum, at least in the areas tested.

In considering measurement issues and their consequences under Model 1, one consequence which will not arise should be noted. Many people presume that the same test which is used to certify pupils' minimal competence will also provide useful diagnostic information to guide remediation. Such is not likely to be the case. While performance on the minimal competency tests will be useful for identifying pupils in jeopardy of being denied a high school diploma, it is not to be expected that a general certification test will provide diagnostic information specific enough to guide remediation. Additional diagnosis will be needed over and above that general information provided by the state tests. It is most reasonable that this diagnosis take place at the district level.

A final consequence of measurement issues in Model 1 is similar to a point made in the preceding section on Standards. The availability of state-wide test performance will lead inevitably to inter-district comparisons of performance. Such comparisons, in turn, can engender a series of pressures on local districts resulting from or accompanied by misuse of the test data (both politically and educationally).

Legal Issues

Legal challenges to minimal competency testing programs will be triggered by arguments of injury, primarily to students who may be denied a diploma. While denying a diploma to a student may not in itself constitute legal injury, the concern in minimal competency programs will likely center on denial of a diploma on the basis of a given set of tests and a given certification standard. Questions of racial-ethnic fairness, validity, timing and the like which will form the basis for litigation are less likely to be directed towards the right of a state to implement a minimal competency program than towards the specific features, tests, standards, and sanctions associated with the particular program implemented.

In one sense, then, many of the consequences of Model 1 we have discussed in previous sections could form a basis for legal action, in particular racial-ethnic bias, test validity, the mesh between tests and instruction, failure to provide remediation, inadequacies of test administration, etc. The central problem in Model 1 is that it is state-wide and thus will tend to overlook, overgeneralize or fail to consider idiosyncracies in local districts, from which legal challenges are most likely to flow. It is for this reason that the need for monitoring agencies to insure local compliance have been stressed throughout this discussion.

Moreover, it is not clear who will bear the brunt of legal challenges - the state, which mandates the program and controls its features; the local district, which presumably has some control over instruction; or individual teachers

who may be faulted for ignoring particular competencies. It is advisable to examine the experiences of other states, Florida and Oregon in particular, to determine the most likely sources of legal challenge and the bases of these challenges. Further, it will be wise to consider specific methods for overcoming the bases for legal challenge by insuring adequate publicity and phase-in periods, validating tests and standards prior to their application, and so on.

Costs

On initial inspection, Model 1 appears to provide a number of economies of scale which may make it an efficient, cost effective implementation strategy to follow. In some respects this is true. Economies of scale result from: the need to construct only one set of competency tests for use state-wide, the need to compare performance against a single standard, the likelihood of centrallized scoring, and the need for a single, state-wide administrative agency. Moreover, at the state level, it may be that particular hardware or distribution mechanisms already exist, thereby removing the need for additional expenditures to purchase or establish these. If we compare our estimated cost of about $45,000 to build a single, state-wide competency test to the cost for each district to build its own test, some notion of the benefit of potential economies of scale is evident.

While potential economies of scale and the reallocation of existing resources may appear to be appealing aspects of Model 1, there likely will be some new or start up costs associated with the implementation of a state-wide minimal competency program. Initial publicity, explanatory materials, and test

validation costs will be required. Of course, the cost-efficient nature of conducting these activities at a state level instead of reproducing them in each local district is still a reality in Model 1.

There are, however, two additional costs associated with a state-wide minimal competency program which must be noted. For the most part, these represent "add on" costs. The need for these two types of expenditures flows directly from the nature of minimal competency programs and their distinctness from other state-wide testing programs. In minimal competency programs, the performance of each individual pupil is of major concern because decisions about granting a high school diploma are made on a pupil by pupil basis. We may contrast this approach with, say, a state-wide assessment program in which decisions about individual pupils are not of interest and data are aggregated across pupil samples. The sanctions and individual pupil consequences which are a part of minimal competency programs, but not other, group oriented state-wide testing programs, will require additional costs.

Under a state mandated, state controlled program, it would seem that the local school districts would have the strongest claim for additional state aid to support local remedial and extraordinary instructional costs. Depending upon the level at which standards of minimal competency are set - and likely varying from district to district - remedial costs may be large. Remedial costs may be pushed higher because, under the minimal competency program, tests are not administered until grade 8 at the earliest. This may not be the most efficient time to initiate remediation and, to do so, may mean a transferral of existing elementary remedial teachers to the junior high or the

"add on" of new junior high remedial services. Funding for earlier diagnosis and remediation, which might have as its goal the correction of pupils' learning weaknesses prior to grade 8 and minimal competency testing, would likely be the local district's responsibility. It is difficult to imagine remedial costs not rising under a testing system which puts a premium on individual pupil performance.

Given the consequences of testing - non-receipt of a high school diploma and its many ramifications - it is evident that a state-wide program such as that envisioned under Model 1, will require extensive monitoring and policing functions. We have described the need for standardizing test procedures, the necessity of insuring that remedial services are provided at the district level, and the importance of being certain that districts are providing pupils an opportunity to learn the state-wide competencies. The administering agency under Model 1 cannot leave these issues to chance, but must implement procedures for insuring district conformity. The need for "inspectors" or mechanisms for certifying district cooperation will be pressing. Note again - and without belaboring an already oft-made point - that the reasons such monitoring activities are necessary is because of the consequences associated with each pupils' test performance. Contrast the consequences of minimal competency testing and state-wide assessment on the individual pupil and the need for monitoring activities should be clear.

Finally, many states, in their rush to embrace minimal competency programs, have dismantled their state-wide assessment bureaus and reallocated

personnel to the minimal competency program. While such a procedure may be cost-effective in the short run, dismantling assessment bureaus can prove costly, when, five years from now, someone wishes to know the effect of minimal competency testing on wider curriculum areas not covered by the competency tests. With no state wide assessment bureau, answering this important question will be difficult and expensive. Consequently, one should think through the benefits and liabilities of dismantling and reallocating existing agencies to a minimal competency program in an effort to be cost-effective.

Model 2: State provided options for competencies, testing procedures and standards from which local districts select

In Model 2, a state or state-sanctioned agency provides a limited set of program options from which each local district may choose in accordance with its own needs and emphases. In many respects, Model 2 is similar - and presents similar consequences - to Model 1, except in Model 2 there are a series of state endorsed programs from which districts can select.

Model 2 has not been implemented in any states and thus is the least documented of the models presented in this report. Due to this fact, there are many unknowns associated with this model and, consequently, more background and prerequisite information needs to be obtained about Model 2 than about the other models presented. Obviously, the model is dependent ultimately upon the nature of the options it makes available to local districts.

If the options are too narrow and homogeneous, there may be little difference between Model 2 and Model 1. If, one the other hand, the set of options provided is very diverse and non-overlapping, the cost and administrative efficiency as well as the programs' credibility with the general public may suffer. What is needed, therefore, and what the viability of Model 2 hinges upon, is information about the current emphases in reading, math, and communications curricula across the state. Without such information any attempt to define sets of options will be arbitrary and may not accommodate local district needs. Moreover, to adopt Model 2 without the necessary prior information about current curricular emphases may result in implementing a minimal competency program which is expensive, administratively unwieldy and which does not truly overcome the problems of local control inherent in Model 1.

The key to Model 2, then, is a knowledge of what currently is emphasized in school districts across the state. Too much local diversity would argue against Model 2 and in favor of Model 1 or 3, since many of the advantages of Model 2 are dependent upon the existence of a limited set of aims and emphases across the state. For example, if there is great diversity in basic skills instruction across districts and the number of options which would have to be made available to satisfy district needs is large (greater than 4 or 5), this in turn would necessitate large test construction and administration costs. Other problems such as program credibility, reciprocity among districts, and legal challenges would likely result from a large number of options from which districts could select.

Therefore, the first step which must be taken prior to adopting Model 2 is an exploration of basic skills instruction as currently carried out in the state. This exploration may take a number of forms, but at the very least and as a manageable first step, it ought to survey the curricula currently in use across the State of Massachusetts. It should survey these to determine first, the degree of diversity in curricula, and second, the overlap among the emphases in these curricula. Such a survey will accomplish two ends: provide a rough estimate of the length of lists of options the state might provide for local districts and suggest how varied testing procedures and standards might have to be to accommodate local diversity. For example, if it were found that the majority of curricula reviewed were similar in regard to their emphases, topics, skills, etc., Model 2 would appear to warrant further consideration as an approach to minimal competency testing. A simple listing of topics will not suffice to identify common competencies stressed in the tests or curricula in use.

For simplicity, consider three possible outcomes to such a survey. In the first case, it may be found that there is very little diversity in basic skills instruction in the areas of reading, communications and math across local districts. In such an instance, Model 2 may not be called for and Model 1 may suffice for implementing minimal competency testing. In the second case, it may be found that emphases vary widely across districts, in which instance Model 2 may be less viable and credible in meeting district needs than Model 3. Finally, it may be found that despite differences across

local districts, there is a limited number of curricular emphases under which virtually all districts can be placed. It is in this third condition that Model 2 appears to be most viable and defensible. There might even be a core set of skills which are emphasized in all districts, with local diversity being a result of particular sets of additional skills tied to this core. From the point of view of implementing Model 2, this situation would be optimum, since it would simplify test construction, scoring, and administration as well as standard setting.

Assuming that a survey showed that there was a sufficiently limited set of district emphases to justify further consideration of Model 2, we may consider some of the administration and resource concerns posed by Model 2. It is probably most reasonable to think about Model 2 and its sets of options as a series of mini Model 1's administered by a state endorsed agency. Thus, instead of the state implementing a single, state-wide minimal competency program as described in Model 1, we may conceive of Model 2 as consisting of the state running three or four programs, with different districts subscribing to different models. Since Model 2 is based upon state endorsed options, including competencies, tests and standards - rather than district specified options as in Model 3 - many of the concerns and considerations about what agency should administer Model 2 are similar to those described in Model 1. The primary difference between Models 1 and 2 is that the administration of Model 2 is likely to be more complex and expensive, since it is necessary to duplicate all the tasks required under Model 1 a number of times

in accordance with the number of options made available for district selection. Thus, multiple tests will be needed under Model 2; greater publicity and explanatory materials carefully outlining differences between options will be required; standardization of a number of testing procedures may be required; etc. Overall, greater resources, both financial and personnel, will be necessary to implement Model 2 than Model 1.

As far as tying testing procedures and standards to lists of competencies is concerned, there are two options available. One is to allow local districts complete discretion in matching competencies to testing procedures and standards (e.g., select one from Column A, one from Column B and one from Column C). There are a number of problems with this option. First, it presumes a level of expertise regarding test validity and standards which may not exist at the local level. It is important that the selected competencies, testing procedures and standards be logically and statistically interconnected to minimize educational, remedial, cost and legal difficulties. Second, the option of "random" selection from three separate lists of competencies, test procedures and standards may encourage districts to select the "easiest" or most passable tests and standards. Third, allowing districts to select options from three separate lists compounds the problems of program administration enormously. Assume, for example that there are three options available within each of the areas of competencies, testing procedures, and standards. Complete freedom in selecting from each list of options can lead to 27 (3^3) separate programs going on at the local district level. The administrative problems associated with a single agency trying to control 27 distinct minimal competency programs are enormous.

In light of these difficulties, it seems more reasonable for the administering agency to permit districts to select a set of competencies which reflect their own needs and emphases and then to mandate testing procedures and standards to accompany the selected competencies. This approach would solve most of the problems listed above by providing greater cost and quality control at the state level and by overcoming some of the program credibility difficulties associated with a smorgasbord selection from three separate lists. Under such an approach, psychometric issues such as test validity, test reliability and standards within options could be attacked in advance at the state level, where the greatest expertise is likely to be available. One of the administrative problems associated with Model 2 is insuring that the various testing procedures and standards within options are comparable, so as to avoid some options being "easier" or less demanding than others. While it is difficult to envision exact test and standard equating procedures across all options under Model 2, it probably will be possible to obtain approximate equivalence across options so that the advantages associated with selecting an "easier" option will be offset by the disadvantages of moving away from already existing local curriculum emphases. Certainly obtaining rough equivalence between options will be easier under the "package" approach of tying testing procedures and standards to particular sets of competencies than under the smorgasbord approach which permits a large number of distinct programs at the local level. The psychometric properties of the testing procedures and the comparability of standards under the various options provided under Model 2

are problematical and will require empirical investigation prior to settling upon particular procedures and options.

Administratively, then, Model 2 presents two general problem areas not encountered in Model 1. First, more preimplementation information and evidence must be collected to determine the feasibility of the model: a survey of current district emphases in basic skills, test and standard equating studies, and so forth. Second, Model 2, when conceived of as a series of mini Model 1's, represents greater administrative complexity and costs than a single state-wide minimal competency program. Assuming, however, that the uncertainties and difficulties of Model 2 can be overcome, we may consider the likely consequences of the model in a number of areas. For simplicity, and in conjunction with the prior discussion, we shall make the following assumptions about the implementation of Model 2: 1) three or four sets of distinct local basic skills emphases have been identified and that these emphases encompass virtually all local districts; 2) districts select competencies and are then assigned state-endorsed testing procedures and standards; 3) the state agency in charge of the minimal competency program has carried out empirical studies related to the testing procedures and standards available under each option.

State versus local Control

An obvious consequence of Model 2 under the above assumptions is that it simultaneously provides an opportunity for local districts to select a minimal

competency program in line with their own current emphases while at the same time maintains a state presence in the certification process. Local control and diversity may be accommodated in this model, while program credibility can be provided by the fact that the model is under ultimate state control. Moreover, the technical expertise required to construct and validate tests as well as determine the comparability of tests and standards across options is placed at the state level where presumably there are the greater expertise and resources to carry out these tasks.

The key to the consequences of a potential state versus local control debate in Model 2 is, of course, the representativeness of the options provided. If the options provided are arbitrary, or reflect the emphases in only half the state's school districts, implementation of Model 2 is not likely to diffuse a state-local control debate, since for a large proportion of districts, the state will be imposing options much like in Model 1.

Impact on Curriculum

Since Model 2 represents an attempt to respond to local district needs and instructional emphases within the context of a state-controlled minimal competency program, it is to be expected that the impact on local curricula will be less under Model 2 than under Model 1. Since districts will be able to select sets of competencies most in line with their current emphases, the need to revamp or reorient local curricula should not be great. Theoretically, therefore, there will be less demand for districts to gear up to a single set

of state-mandated competencies which may be disjoint from current emphasis. Moreover, less teacher retraining will be required. As noted previously, regardless of the minimal competency model adopted, the greatest impact on the curriculum in each local district is likely to arise in the area of communications skills, since overall, these are probably the least emphasized skills encompassed by the presently envisioned minimal competency areas.

Model 2 as described puts increased pressure on districts to do well what they currently are seeking to do - insofar as graduation will be dependent upon satisfactory pupil performance of the locally stressed competencies. One potential, more long range consequence associated with Model 2, is the possibility that debate within districts will arise regarding which set of state approved competencies a district should be seeking to foster. That is, the availability of multiple options may lead logically to comparisons of the desirability of the varied options, quite independently of the mesh between current district emphases and the options. Within communities, debate over the "appropriate" competencies for pupils may arise. If such debate leads to scrapping the present curriculum in favor of an option which is judged to be more desirable or appropriate, then the impact of this change on local curricula will be great. Overall, however, barring such long range changes, the impact on the curriculum at the local level under Model 2 is not likely to be substantial - except in areas not presently emphasized such as communications skills.

Remediation

As noted previously, it is most likely that specific remedial activities will be carried out at the local district level, regardless of the model of minimal competency adopted. In the initial years of a Model 2 minimal competency program, there are some potentially beneficial consequences which could ensue vis a vis remedial activities. Given that the option selected by a local district under Model 2 is similar to its current instructional emphases, established remedial services, also in line with these emphases, should already be in existence. While Model 2 might require some shifting of these services from the elementary to higher grade levels in line with the proposed timing of minimal competency testing, it seems reasonable to assume that many remedial problems could be anticipated and corrected in earlier grades as part of the ongoing instructional program. That is, with a correspondence between district emphases and the competencies required for MCT, early diagnosis and correction of learning problems could take place, thereby anticipating at early grades many of the problems which can occur in later grades. Further, it would be in local districts' interests to diagnose and remediate early, insofar as remedial costs for gearing up a junior or senior high remedial program could be diminished.

It is important to note that the scenario just described is likely to be true only in the early years of Model 2. Under another model, say Model 1, more extensive local curriculum revision may be required, so that the advantages of early diagnosis and remediation would have to await the

installation and integration of new curricula which responded to the state-wide competencies. Once such curricula were installed, however, districts under Model 1 would be in a position to carry out early diagnosis and remediation. The advantage in Model 2 is simply that the initial curriculum revision and installation phase may not be required.

Overall, as in all models for minimal competency testing, it is likely that the need for remedial services will increase over present needs. This increase will be due to the use of an external certifying exam as a prerequisite to high school graduation. Districts - and the state - will have to monitor pupil progress more carefully and provide remedial services where necessary, if only to defend the legal validity of denying some pupils a diploma. As noted, Model 2, by gearing the minimal competency program to existing local district emphases, presents certain short range advantages vis a vis remediation.

Standards

In Model 2, there likely would be different standards associated with the different district options, since this procedure is more defensible than building exactly comparable tests across options. Across options, there may be difficulties in precisely equating standards, since the technology for such equating leaves much to be desired. However, it seems reasonable that a rough equating can be accomplished. At any rate, it will be easier to justify the equivalence of three or four standards as in Model 2, than literally hundreds of standards as in Model 3.

Regardless of the nature of the minimal competency program implemented, there will be certain consequences associated with standards which will arise. Primary among these consequences are challenges to the standards on the basis of racial or ethnic bias. Similarly, regardless of the model implemented, there likely will be inter-district comparisons of passing and failing rates. In seeking to respond to local district needs and emphases, Model 2 potentially exacerbates these consequences and raises additional ones.

For example, questions of racial-ethnic bias are likely to result in the context of three or four options under Model 2 as opposed to one option under Model 1. Moreover, the existence of multiple options will invite comparisons between racial-ethnic performance across options, and unless standard equating is successful, lead to debate about the bias of various options towards racial and ethnic groups. The existence of options as in Model 2 also invites inter-district comparisons both between and within options. Within model comparisons of performance will be similar to those described under Model 1, except that comparisons will be made within three or four different options. The potential for between option comparisons is more dangerous however, since these may be done in an uninformed way (e.g., ignoring substantive differences between options) or in a way that puts pressure on a community to drop its current emphases and replace these with the emphases of a different option. In general, the potential for inter-district comparisons and the consequences - positive and negative - which result from such comparisons is greater under Model 2, with its multiple options, than under other models such as Model 1. Finally,

the existence of varied options within the minimal competency program gives rise to concern about inter-district reciprocity vis a vis the options. Two issues are of concern with regard to reciprocity. First is the equivalence of the standards set under each option. Second is the similarity of the particular competencies required by different districts. For example, if a student leaves district A at the 9th grade and moves to district B which follows a different minimal competency option, which option (i.e., which set of competencies) is the student to be held responsible for? Note then, that the reciprocity issue under Model 2 - and Model 3 for that matter - involves not only the comparability of standards, but also the nature of the competencies required.

Measurement Issues

The measurement issues and consequences associated with Model 2 are similar to those described under Model 1, except that they are compounded by the diversity implicit in local options rather than a single state-wide program. Thus, the standardization of test administration timing and procedures across various options will be necessary in Model 2, just as in Model 1. This task may be accomplished by standardizing within each option or by standardizing timing and procedures across options in the form of a state-wide minimal competency testing period. Similarly, the test security problem and the need to revise tests yearly will occur within Model 2, but the problem will be exacerbated by the need for multiple tests in line with the various options. Test validity vis a vis the competencies and the instruction provided

to students will need to be determined and monitored as in Model 1. This monitoring activity may be more straight forward in Model 2 because districts presumably will be following options in line with their current curricular emphases. Virtually all of the measurement consequences associated with Model 1 will be required in Model 2, with the added consequence that the economy of scale associated with Model 1 will be diminished because of various options within Model 2.

Legal Issues

As in the case of measurement issues, many of the legal issues associated with Model 2 are similar to those likely to occur in Model 1. In fact, many of the legal consequences of minimal competency programs tied to high school graduation are endemic to any implementation model. Thus, questions of test or standard bias are likely to arise in any program, as are questions of test validity, including concern over direct versus indirect measurement and timed versus untimed tasks.

There are two potential legal consequences associated with Model 2, however, that are not associated with Model 1. On the one hand, the existence of a number of options within the Model inevitably leads to questions about the comparability and fairness of the tests and standards across the various options. If, for example, one option could be shown to have more difficult tests than another, could legal questions result based on comparability across options and the fairness of denying a higher proportion of pupils in the more difficult option a high school diploma than in the easier option? It is to forestall

problems such as these that the need to pretest and validate tests and standards was noted above. Note, however, that in Model 2, the state controls test development and standard setting within the options and thus presumably can bring more expertise and resources to bear in overcoming such potential problems than if each district had responsibility for determining its own standards and testing procedures. Moreover, could racial-ethnic groups within a community bring suit against the community because it had adopted an option they felt was unfair or irrelevant to minority pupils? Thus, the existence of a number of options within a minimal competency testing program may increase legal challenges to the testing procedures and standards.

On the other hand, the fact that local districts could select options in light of their own needs and emphases likely would reduce the legal challenges resulting from arguments that the schools didn't teach the required competencies as part of their instructional program. Similarly, claims that remedial activities were not provided or were not relevant to the tested competencies would likely be less prevalent under Model 2. Overall, then, there are some potential advantages and some potential disadvantages of Model 2 relative to Models 1 or 3 in the legal domain.

Costs

By centralizing overall administration of Model 2 at the state level economies of scale are capitalized upon. However, it is obvious that the increased diversity of Model 2 in comparison to Model 1 tends to reduce the benefits of ecomonies

of scale. The need to police and administer three or four options, instead of just one; to construct and validate many sets of competency tests, instead of just one; to administer and score many different tests, instead of just a few, all contribute to increased costs under Model 2. More personnel are needed for administration, for test construction, for monitoring the activities of local districts, and so on.

If there is any potential financial advantage of Model 2 over a centralized, single state-wide minimal competency program it rests in either of two areas. Because the local district has options to choose from in selecting its minimal competency program, it seems reasonable to expect that the local district could be required to assume greater responsibility towards providing funding for the program. Moreover, it would seem that the local district could make its greatest cost contribution in the areas of remedial expenses and test administration and scoring costs. While some of these expenses - as well as expenses to be incurred at the state level - might be met by diverting existing testing and remediation funds to the minimal competency program, there are two issues, noted previously, which argue that most of the expense of minimal competency testing programs will be "add on" costs. First, to divert funding from current pupil and state-wide assessment programs is to lose the future potential to examine the effect of the minimal competency program on wider areas of the curriculum. In order to assess the impact of the program on pupils performance in other curriculum areas, it will be important to maintain current assessment operations. Moreover, the

information about education in the state which is provided by state-wide assessment and which is different from the information provided by MCT, will also be lost if the current assessment program and staff is diverted to other tasks. Second, and perhaps more importantly, minimal competency testing requires a different, more detailed and validated testing technology than general pupil or state-wide assessments, since the consequences associated with minimal competency test performance are considerably graver than those associated with typical assessment programs. Additional resources for constructing and validating tests, record keeping, validating standards, etc. will be required to make the minimal competency program legally, educationally and ethically defensible. It is unlikely that these additional requirements for minimal competency testing can be accomplished without substantial "add on" costs.

Model 3A: Local responsibility for setting competencies, testing procedures and setting standards subject to state approval

Model 3B: State responsibility for setting competencies, local responsibility for testing procedures and setting standards subject to state approval

In selecting the models to be examined in further detail, the policy subcommittee expressed interest in models which allowed each local district (or possibly a group of districts) to develop its (their) own specific minimal competency testing program subject to state approval. Two closely related models were selected. In one model, the local districts handle all three major functions, setting the competencies, developing or selecting the tests,

and setting standards of mastery, with these three functions subject to state approval (Model 3A). In the second model the state sets the competencies and the local districts develop or select the tests and set standards of mastery, with the latter 2 functions being subject to state approval (Model 3B). Since these two models differ only with respect to who sets the competencies, the state or the local districts with state approval, they are treated together in this section.

Let us first examine and contrast Models 3A and 3B on the one area in which they differ, namely who sets the competencies which are to be tested. In Model 3A, the local districts bear the responsibility of drawing up a list of competencies in each area to be assessed (i.e., math and communication). In keeping with the suggestion that a broad representation of groups and interests be involved in setting the competencies, Model 3A affords the opportunity of involving many more people than does Model 3B. The sheer number of people involved is not as important as the fact that local people, close to local needs, are involved. Models 1, 2, and 3B are all lacking in this regard. Each district or group of districts would establish committees composed of representatives from business and industry, as well as educators, parents, and citizens. Thus, in Model 3A local people would be actively involved in the development of the local program. This involvement of representatives of local interests should increase the acceptance of the MCT program. It should be noted, however, that Oregon's experience shows that local districts were overwhelmed by the process of setting competencies and eventually asked the state to undertake this task.

The end result, i.e., the lists of competencies, that are derived by different districts under Model 3A may overlap considerably. This overlap could be seen as a disadvantage in terms of massive duplication of effort throughout the Commonwealth. Model 3B circumvents this drawback by centralizing the process of setting competencies. Because of this centralization, a sense of local ownership of the program, at least with respect to the setting of competencies, is lost in Model 3B. Under Model 3B, there may be resentment at the local level because the competencies to be taught were mandated from above. This resentment is likely to be lessened by affording districts latitude in selecting testing procedures and standards. However, it should be noted that a single set of state approved competencies does provide some external credibility to the minimal competency program as well as helping to standardize the meaning of a high school diploma.

A second consideration related to Model 3B is that some local districts have already begun to implement their own minimal competency testing program. It may be difficult to tell these districts that their work was for naught, and that they now have to adopt the state specified competencies. If Model 3B were adopted, some special concessions may have to be made to accommodate those districts which have an existing program. There are problems with making any concessions however, including the danger that the whole effort could be undermined. Such concessions may be viewed by other districts as preferential treatment, and this attitude could result in ill-will toward the MCT

program and even between districts. The trade-offs have to be carefully weighed under Model 3B if districts with existing minimal competency programs are to be dealt with differently from the majority of districts that do not as yet have such programs.

With the state setting competencies in Model 3B, economy of scale is a major advantage. It seems doubtful that any LEA would possess the staff and resources necessary to develop competencies to the same degree as would a state agency. A further advantage to having the state set the competencies is that reciprocity in terms of accepting the certification of competencies between districts within the state would be easier to handle; all students in the state would be taught the same set of competencies. With our highly mobile society, one could argue that reciprocity is almost as much of an issue between states as it is within any one state. Although this may be true, we must confine ourselves at this juncture solely to Massachusetts.

Additional disadvantages associated with the state setting the competencies include the end-product being viewed as an attempt at a state controlled curriculum. Local districts may disagree with the competencies and be reluctant to sacrifice their local control over curriculum in order to teach the state required competencies. There is also an assumption that the state required competencies are accepted by the local districts, since the local districts are responsible for testing the competencies and setting standards for mastery under Model 3B. Local districts could attempt to side-step the

competencies by submitting tests for approval that are suited to their existing curriculum, and not to the state required competencies, in the hope that the state agency will be so overloaded with tests to review that any test will be acceptable. If such a situation is allowed to happen the whole minimal competency testing program could soon become a sham.

When the local districts set their own competencies, the issue of a common state curriculum appears to be circumvented. However, if the state agency has specific competencies in mind and rejects all locally submitted competencies unless they contain the competencies desired by the state, the end result is much the same as the state setting the competencies at the outset. The process, when the local districts set the competencies, forward them for approval and revise until the state approves them, takes much longer and may be more wasteful and frustrating to all involved.

For either of the models being discussed in this section, some state agency would have to review tests, determine whether the tests adequately measure the competencies, and decide on whether the criteria for mastery are appropriate. In the model in which local districts also submit their own list of competencies for state approval, this state agency would also have to decide on the appropriateness of the competencies. Accomplishing these tasks will be difficult. There is so much subjectivity involved in determining the appropriateness of competencies, tests, and standards, that it may be difficult to justify rejecting a locally submitted plan. If a local plan is rejected, an appeal process may have to be available to the local district.

The same problems inherent in the state agency rejecting competencies submitted by local districts exist for the other two components of the model also. The extent to which the state has narrow criteria of acceptance is the extent to which the local districts really only have the appearance of control over their own minimal competency testing program. In effect, actual state control over the three components of the model could be preferable to very rigid, narrow acceptance criteria by the state of locally submitted plans. The issue of what agency at the state level is to oversee the projects is an issue in Models 3A and 3B as it was in the other models. We can not resolve this problem but can merely point out that an existing agency, a totally new agency, or an outside contractor could be employed to handle all or any of the three components of either Model 3A or 3B.

Having discussed these two models in a general way, we will focus in turn on the 7 issues on which all the models presented in this section may be compared: 1) state vs. local control, 2) impact on curriculum, 3) remediation, 4) standards, 5) measurement issues, 6) legal issues, and 7) costs.

State versus Local Control

Aside from the obvious state versus local control difference between Models 3A and 3B in setting the competencies, there is the issue of how much local control either of these two models really offers. At first blush, it would seem that Models 3A and 3B afford the local districts much more control than either Model 1 or Model 2. In fact, Models 1 to 3 appear to lie along a continuum from strong state control to strong local control. As we

have pointed out, however, the stringency of the state approval criteria affects how much control the local districts actually have over their own program. Rigid and narrow criteria by the state for approval of local plans can shift Models 3A and 3B closer to Model 2 or possibly even Model 1. It is difficult to anticipate how much control either Model 3A or 3B will actually give to the local districts. This committee should simply be aware of the fact that the potential is there for the amount of local control in either of these two models to be less than it appears at first glance.

A further issue under Model 3A, and to a lesser extent under Model 3B, is the credibility of these models. It may be that locally developed competencies, tests or standards will be viewed skeptically by the public, whereas a state-run program may be seen as having greater credibility. Again, there are trade-offs that need to be weighed between this credibility issue and giving local districts more say in the structure of their program. Another concern under Models 3A or 3B is whether the local districts have or will have the expertise and resources needed to mount such a program. Particularly, for smaller districts, one may be imposing a financial hardship on the district under Models 3A and 3B.

Impact on Curriculum

Since the competencies required will dictate in large part the impact on curriculum, Models 3A and 3B could differ in their impact. Local districts may set their own competencies so that they reflect the existing curriculum. To the extent that this occurs, little impact on curriculum will result. State

set competencies may have more impact on curriculum, but again this depends on how different the competencies are from what is currently in the curriculum. As stated in the earlier section dealing with impact on curriculum, competencies in listening and speaking may result in greater curriculum change than competencies in math, reading and writing. The former are not currently emphasized in most curricula as heavily as the latter.

Remediation

Under any of the models discussed in this report, there is a need for a remedial component as part of a MCT program. With local districts developing their own program, there may be greater opportunity for them to obtain diagnostic information on their students and thus, enhance their remediation efforts. For example, a local district involved in developing their own minimal competency tests may find that much of their effort can also be applied to developing diagnostic tests to be given in the early grades. As a result, there may be greater cohesiveness or integration between diagnosis/remediation and the competencies required by the minimal competency tests. If local districts have access to staff skilled in test development, in order to implement a minimal competency testing program, this staff may also be able to develop a whole series of instruments which provide diagnostic information to teachers on an ongoing basis. In many cases, teachers will be able to provide pertinent diagnostic information based on their classroom experiences with pupils. Educational benefits to the students and teachers, other than maximizing

the student's chances of passing the minimal competency tests, would accrue. In smaller districts, where it may not be feasible to have one or more people assigned solely to developing a minimal competency program, consultants could be hired or collaboratives of 4 or 5 districts could share a person skilled in test development.

Standards

As we have pointed out, the problems associated with establishing standards or cut-scores are many and complex. There are not only statistical issues, but also political and economic ones, which impinge on setting standards. Again, someone with expertise in this area would have to be available to the local districts, if they are expected to set their own standards for passing the minimal competency tests.

More importantly perhaps is the issue of comparability of standards across districts. Even though the state would have to approve each district's standards, there would be no way to judge whether standards were uniform from district to district just from examining a test and the proposed cut-off score. The only way to insure that standards are uniform across districts is empirically. Some students in district A would take district B's tests in addition to their own district's tests and vice versa. Comparisons between the two district's standards could then be made on the basis on how similarly the two tests categorize the students. Of course, this procedure becomes muddied if the competencies for the two districts are different, since district B's tests may not be appropriate for district A's students and vice versa.

Further, although comparisons between districts may not be justifiable, because of different competencies, tests, composition of the communities, etc., such comparisons will probably occur. Issues such as, "District A requires only 70% correct for passing their tests while District B requires 80%" are sure to arise with the unwarranted conclusion being that District A is easier than District B. Problems may result from different districts having, what appear to be on the surface at least, different standards.

Finally, the issue of reciprocity between districts is particularly acute under Model 3A. A student who transfers from one district to another in grade 8, for example, may be tested on competencies very different from the ones he was taught. Concessions may have to be made under this model for students falling into this category.

Measurement Issues

Local districts could either develop their own instruments or contract out this task. In either case, an added economic burden would be placed on the local school budget, but more will be said about this point in the cost section. The development of instruments includes a validation process. Depending on how secure the tests are kept, test development (including validation) is an ongoing process. It is unclear to us if this validation process would have to take exactly the same form for Model 3A or 3B as it would for a state-wide test, for example. It may be easier for a local district to demonstrate that its tests accurately reflect its curricular emphases than it would be for the state to demonstrate that its tests accurately reflect the curricular emphases in all local districts. Evidence that

the tests adequately measured what they were intended to measure would have to be obtained, but again it is unclear if this evidence would have to be as strong here as it would for a state test.

Issues such as differential failure rates in urban as opposed to suburban districts may be partially circumvented when each district has its own program. The question would arise, however, as to the comparability of different instruments across districts. It may be that some districts will get a reputation for being "easy" and others "hard." A possible implication of this outcome may be migration to the easy districts by families whose children may be denied a diploma in a hard district. The task of assuring comparability across various district's tests is very difficult. As with equating cut-scores, the state approval process could not guarantee equivalence by merely inspecting the various tests. Empirical studies may help in the equating process, but the issue of comparability of tests across districts is deserving of very close attention in Model 3A or 3B.

Another aspect of comparability between districts is the timing of test administration. Will all districts have to test their students beginning at the same grade or can one district certify students beginning at grade 8 and others beginning at grade 10? Will administration dates have to be the same across all districts in the state or can some districts test on October 15, others on January 15, and still others on May 30? If each district controls its own program, statewide standardization of testing time and procedures would probably not be necessary as in previously discussed models.

Legal Issues

The legal challenges that may arise are intertwined with many of the prior issues just discussed. Local districts may be challenged by an individual student who is denied a diploma or by a class of people who question the whole testing program. It would seem that the potential for legal challenges may be lessened when each local district has its own program. This may be an instance where economy of scale works in reverse, however, it is difficult to be sure on this point. Local districts currently make decisions that affect individual students and very few of these decisions are subjected to legal challenge. This is not to imply that legal challenges will not occur under Models 3A or 3B, but simply that the frequency of their occurrence may be less than under a state run program.

If legal challenges do arise under Models 3A or 3B, what form could they take? Locally developed instruments may be technically weakest. Thus, a local program may be challenged on technical, psychometric aspects of their tests. Some have questioned whether this is a real possibility, since local testing is already an integral part of local education. Teacher-made tests are not generally technically sound, yet such tests are not challenged. There may be some weight to this argument, but when a single test, with the power to deny a high school diploma, is used, the situation is more focused and is quite different from the teacher-made test situation.

Other challenges may arise from differential failure rates by various socio-economic, ethnic, or racial groups. This type of challenge is more apt to arise in districts where there is a mix of these groups than in an upper middle class, white suburban district.

Local districts will need to keep accurate records on each student, specifying when a particular competency was taught and attained or remediated. Local districts will also have to monitor staff to insure that each teacher is teaching the competencies and keeping accurate records on each of his/her student's progress toward attaining these competencies. Only with accurate record-keeping can legal challenges be withstood.

Again, it is difficult to anticipate all possible grounds for legal challenges and the extent to which legal challenges will occur under Models 3A or 3B, but the potential for such challenges certainly exists.

Costs

It would appear that the more responsibility each local district has for developing its own program, the greater the monetary costs will be. Most local districts are not currently staffed with people capable of developing instrumentation, carrying out validity studies, etc. Yet people with this expertise are necessary if a strong and defensible minimal competency program is desired. Consultants or contractors could be brought in to manage the program, but cost is still a factor here. One could argue that the additional cost for a MCT program is well-spent, since it provides some evidence as to whether the remainder of the educational budget, certainly the preponderance of the money spent on education, is at least producing students who have minimal competencies. Testing and evaluation are often viewed as frills and expendable, when in fact they can be an important component in justifying the rest of the budget or in deciding how best to allocate the rest of the budget.

Nevertheless, the overall costs incurred under Model 3A or 3B will be higher than under either Model 1 or Model 2. If there is no local support for Model 1 or 2, however, money spent on them may be money totally wasted, since the program may never become operational. In this case, the additional monetary costs incurred under Model 3A or 3B are certainly justifiable.

As with all the models, development costs, legal costs, record-keeping costs, remediation costs, additional staffing costs, etc. must be absorbed under Model 3A and 3B; the local districts may be hard pressed to obtain state monies for absorbing these costs, since the programs may be viewed as local programs. If the state is running the program or has a great deal of control over the program, as in Models 1 or 2, local districts may fare better in terms of the state bearing the bulk of the costs for the program than they would under Models 3A or 3B. We are not merely talking here about state funds for the program per se. The related expenses that may be incurrred from a minimal competency program, such as increased record-keeping and remedial services, may have a better chance of receiving state funds, if the state has a more active role than it may have under Models 3A or 3B. This is not to say that no state funds would be available for these models, but that the chances for a greater amount of state funding seem better under one of the other models. At the very least, local districts would have a better case for requesting or demanding state funds when the state controls the program in large part.

Summary

At this point, we had intended to summarize the preceding comparisons of models by providing a table in which the models were rated relative to each other in terms of the following 7 issues: 1. strong state control, 2. large impact on curriculum, 3. increased remediation services, 4. difficulty in setting standards, 5. many measurement problems, 6. many legal challenges, and 7. high costs. A table was laid out and a rating scheme devised. The authors then independently attempted to complete the table using the rating scheme shown at the bottom of the table.

This exercise was profitable in many respects. First, we found that we could not agree on the ratings for the seven categories. As a consequence, we are unable to present a completed comparative table in this report. Second, we realize that we could not agree because, as we have been suggesting throughout this report, the issues are complex and interact with one another. When one of us rated the models relative to the difficulty in setting standards, for example, he tended to key in on a set of issues very different from the others.

Our attempt at building a table that compared the models on 7 issues dramatized the complexity of the issues and the futility in attempting such a simple, one dimensional summary. Since the exercise was so helpful to us, we have included the table, without ratings, so that you may attempt the exercise. If you are able to confidently make all ratings, you got further than any of us.

Table 1

Comparison of the 3 Models
Across 7 Issues

Issues	Models			
	Model 1 State sets competencies, tests, and standards	Model 2 Local districts select from state packages	Model 3A Local districts set competencies, tests and standards subject to state approval	Model 3B State sets competencies, local set tests and standards subject to state approval
1. Strong state control				
2. Large impact on curriculum				
3. Increased remediation				
4. Difficulty in setting standards				
5. Many measurement problems				
6. Many legal challenges				
7. High costs				

Note: Use "+" to indicate that the model embodies a great deal of the particular issue, "o" to indicate a moderate amount, and "-" to indicate very little. All ratings for an issue should be made relative to the other models.

PART IV: RECOMMENDATIONS FOR IMPLEMENTATION

In this final section, we make some general recommendations on how we feel Massachusetts may best proceed in making a final decision about the adoption of a MCT program tied to high school graduation. We do not recommend a particular model, nor will we recommend whether a MCT program of some sort should or should not be adopted. Our recommendations take the form of more general suggestions and questions that we feel should be answered prior to making any decision about the adoption of a MCT program tied to high school graduation.

First, we feel that all groups who have been working on the issues and problems associated with minimal competency programs in the Commonwealth are to be commended for their careful approach to this issue. Unlike so many other states, Massachusetts has avoided the impulse to institute a program without careful consideration of all the ramifications of such a decision. Minimal competency programs are the current "in" movement in education, and it is difficult not to be swept up in the movement. Massachusetts has succeeded, thus far, in carefully considering the advisability of a minimal competency program and in avoiding being swept along by the tide of opinion.

The problems that other states have encountered because they rushed to adopt MCT have been avoided thus far in Massachusetts. The opportunity to learn from the problems others have encountered is a felicitous and rare one for Massachusetts. If anything can be learned from previous "innovations" in education, it is that they are more complex than initially imagined and not always as beneficial as was first believed. The pressures are great to adopt some sort of MCT program, however,

Massachusetts has taken the right course in their deliberate and careful consideration of the issues.

Our first, and strongest recommendation then, is to continue on the present course of careful consideration and planning. It would be unwise to let the efforts thus far go for nought because of pressure to enact a program immediately. There are many issues raised in this document for which there are no easy answers. As far as possible, answers to many of these issues should be obtained prior to making the final decision on whether or not to implement a MCT program tied to high school graduation. In what follows, we have delineated what we see as the issues that need further study and offer suggestions on how to obtain more information on these issues.

Of utmost importance is that various publics, including the state legislature, parents, and school people, be educated about what a MCT program tied to high school graduation really means in all of its ramification. All too often we assume that there is an understanding across various publics in regard to an issue, when in fact many envision something very different from what the committee might have intended. It is our impression that many publics view a minimal competency program as simply an extension of some statewide assessment program. However, as noted, the two are very different entities involving very different procedures, needs and consequences. A minimal competency program has direct impact on every individual student. A strong sanction is applied for poor performance. Decisions - often of life long importance - are made about individuals in minimal competency testing programs. This is not the case in statewide assessments.

The public needs to be made aware of these facts and of the potential problems that can arise. Various surveys, among these one conducted by this committee, conclude that a vast majority of most publics currently favor a MCT program tied to high school graduation. A valid concern is whether this judgement is in many cases an informed one. We therefore recommend that an intensive effort be made towards alerting various publics to the meaning, problems, and implications of adopting a minimal competency program tied to high school graduation. This effort implies, of course, that there currently is some agreement on the need and ultimate purpose of a minimal competency testing program. If this effort is successful, some of the pressure to act quickly may be reduced, and a more informed portrayal of the sentiments of various publics can be obtained.

The general issue of phase-in time or time between deciding that Massachusetts will have a program and when the program will go into effect, involves a series of questions. First, "How much time is needed to permit local districts to gear up their curriculum offerings in line with the desired minimal competencies?" As we have stated in Part II, the extent to which the desired competencies are currently part of the existing curriculum relates to this question. We recommend, therefore, that information be obtained regarding the minimal competencies that currently are common to and those that are unique to curriculum throughout the state. A well planned survey of teachers' stratified by local districts, which asks them how much emphasis they place on each of the minimal competencies being considered would accomplish this purpose. From this information, determination could be made as to the overlap of curriculum emphases across local districts and the overlap of each district's

curriculum with the proposed minimal competencies. Such a survey would also serve the purpose of alerting teachers to the types of competencies that are being considered as requirements for certification. The results from this survey would allow one to determine how much curricular revision would be needed so that realistic estimates of phase-in time could be made. At the same time LEA administrators can be polled for their opinions about the advisability of a minimal competency program, the preferred implementation strategies, and the concerns associated with such a program. This information can all be used in assessing the most desirable phase-in period.

Also related to the amount of phase-in time is the issue of test development. Questions such as, "How much test development and validation are required before minimal competency tests can be regarded as valid measures?" need to be answered. Test publishers and consultants should be asked to provide time estimates on how long it would take them to develop, and validate, various types of instrumentation. A study of the costs of setting up a test development bureau to service a MCT program should also be conducted.

The preceding information can also be used to guide the committees who are to finalize the competencies. We recommend that these committees represent a broad cross-section of the publics in Massachusetts. Many such committees, comprised of a manageable number of people, may have to be convened with representatives from these committees then meeting to finalize the desired competencies. Unless this approach is taken, there may be little support for even the best of programs.

More counsel on how to handle special situations, such as physically handicapped students, mentally retarded students, English as a second language students, "hard core underachievers," etc. should be obtained. There are psychologists and special needs teachers, among others, who should be consulted about issues of MCT with these populations. It is not clear how much phase-in time is needed to meet the special needs of these pupils vis a vis minimal competency testing. A large part of the legal problems emanating from the Florida minimal competency testing programs are tied to inadequate phase-in periods for such special learner groups.

Time will also be required to assign defensible standards to pupils' test performance on the minimal competency tests. The proposed final test instruments should be administered to a try-out sample of students so that baseline data can be obtained and the viability of various standards or cut-scores can be determined. At the same time, information about the differential failure rate by various student populations may be obtained.

With respect to instrument validation, the proposed tests should be administered to various segments of the adult population. For example, groups of college graduates, graduates of high school only, high school dropouts, prisoners, unemployed, blue collar workers, white collar workers, etc. could be tested. A number of questions could be answered by this approach. The primary validity question is how well does the test differentiate between successful and unsuccessful adults, however success is defined in terms of some of the above groups. Another question that could be answered is, "What proportion of the current adult population possesses the minimal competencies we expect of our students?"

The results from testing adult populations could be very helpful in assessing the appropriateness of the criteria or cut-scores also. An additional benefit of testing adult populations is that they will be acutely aware of what a minimal competency test looks like and what competencies are being required.

Most of all, however, we recommend that minimal competency test performance not be tied to high school graduation until the above issues, data and phase-in periods provide satisfactory evidence that it is appropriate and valid to tie test performance to graduation. As noted, the consequences of minimal competency testing programs tied to graduation are grave. The fairness of any program which makes minimal competence a prerequisite to high school graduation is based upon the issues and phase-in questions raised above. Pupils should not be penalized on the basis of faulty procedures or inadequate information, which, with proper reflection and phase-in periods, could have been rectified.

Finally, in considering the models presented in Part III, or any implementation models for that matter, we strongly urge that the fact that there are strengths and weaknesses, pros and cons to every model be kept in mind. The issues affecting a model interact with each other. A strength of one model may impinge on other aspects of the same model. For example, some would view strong local control as a strength of Model 3, but monetary costs go up with this decentralized model. Thus, in deciding on models, it is a constant trade-off of issues or strengths vs. weaknesses of a model.

Further information on the pros and cons of the various models could be obtained by trying out alternative models in various districts. Problems that are

not apparent at this point in time may emerge through this approach. Questions about the viability of various aspects of a model could also be answered through this approach. The effectiveness of alternative implementation models could be compared.

The thrust of our recommendations can be summarized quite simply: Be deliberate in making decisions about the advisability of adopting a minimal competency testing program tied to high school graduation, anticipate the issues involved in such a program and try to obtain as much information about these issues prior to implementing a program. Pilot programs which differ in nature could be implemented and compared. In any event, no program need be viewed as the final one. We can learn from our mistakes and improve on what we have begun. This evolving approach to a competency program seems to make the most sense. The tasks at hand are formidable ones, but they can be accomplished if decisions are made rationally rather than in haste.

BIBLIOGRAPHY

American Psychological Association. Standards for educational and psychological tests. Washington, D.C.: American Psychological Association, 1974.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. Taxonomy of educational objectives. New York: McKay, 1956.

Cicirelli, V.G., et al. The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development. Volume I, Westinghouse Learning Corporation and Ohio University. U.S. Department of Commerce/National Bureau of Standards/Institute for Applied Technology. Distributed by Clearinghouse, Springfield, Va., PB 184 328, 12 June 1969.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. and York, R.L. Equality of educational opportunity. Washington, D.C.: Office of Education, U.S. Department of Health, Education and Welfare, 1966.

Edmonds, R. Developing student competency: Alternative means and attendant problems. Harvard University, 1977 (mimeo).

Foley, J.J. Evaluation of learning in writing. In Bloom, B.S., Hastings, J.T. and Madaus, G.F. (Eds.) Handbook on formative and summative evaluation of student learning. N.Y.: McGraw-Hill, 1971, pp. 767-814.

Glass, G.V. Standards and criteria. Paper #10 in Occasional Paper Series. Kalamazoo, Mich.: Evaluation Center, Western Michigan Univ., 1977.

Glass, G.V. Postscript to "Standards and Criteria." Paper presented at 1977-78 Winter Conference on Measurement and Methodology of the Center for the Study of Evaluation. University of California, Los Angeles, 1978.

Jackson, R. Developing criterion-referenced tests. Princeton, N.J.: ERIC Clearing House on Tests, Measurement and Evaluation, Educational Testing Service, 1970.

Jencks, C., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B. and Michelson, S. Inequality: A reassessment of the effect of family and schooling in America. N.Y.: Basic Books, 1972.

Jencks, C. The wrong answer for schools is: (b) Back to basics. The Washington Post, Sunday, Feb. 19, 1978.

Lerner, B. Equal protection and external screening: Davis, DeFunis and Bakke. Proceedings of the 1977 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service.

McClung, M.S. Competency testing: Potential for discrimination. Clearinghouse Review, August 1977, pp. 439-443.

Madaus, G. F. & Airasian, P.W. Issues in evaluating student outcomes in competency based graduation programs. Journal of Research and Development in Education, Vol. 10, No. 3, 1977, pp. 79-91.

Madaus, G. F. & Rippey, R.M. Zeroing in on the STEP Writing Test: What does it tell a teacher? Journal of Educational Measurement, Vol. 3, No. 1, 1966, pp. 19-25.

Madaus, G. F., Nuttall, R. L., & Wood, E. A causal model analysis of Bloom's taxonomy. American Educational Research Journal, Vol. 10, No. 4, 1973, pp. 253-262.

Tractenberg, P. L. & Jacoby, E. Pupil testing: A legal view. Phi Delta Kappan, Dec. 1977, pp. 249-254.

Turnbull, W.S. Statement before the Subcommittee on Education, Arts, and Humanities. Committee on Human Resources, U.S. Senate, July 27, 1977.